

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 18:29:19 ; Search time 224 Seconds
(without alignments)
3205.659 Million cell updates/sec

Title: US-10-646-396-2
Perfect score: 6619
Sequence: 1 MDKARNCLLQHREALEKDI.....FKTVTVVDNLGILVLOTLE 1248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trenbl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6619	100.0	1248	1 APAF HUMAN	O14727 homo sapien
2	5873.5	88.7	1249	1 APAF MOUSE	O88879 mus musculus
3	5871.5	88.7	1249	1 APAF RAT	Q98pvs rattus norv
4	5863.5	88.6	1249	2 Q8V166	O8vi66 rattus norv
5	4054	61.2	1248	2 O6GNU6	O6gnu6 xenopus lae
6	3683.5	55.7	1261	1 APAF BRARE	Q919h8 brachydanio
7	1066	15.1	258	2 Q80VR5	Q80vr5 mus musculus
8	732	11.1	1227	1 Q8ZOR1	Q8z0r1 anabaena sp
9	728.5	11.0	1258	1 Y800 ANASP	Q8ytc2 anabaena sp
10	693	10.5	1184	2 Q7ND85	O7nd85 gloeobacter
11	691	10.4	1526	1 YK46 ANASP	O8yril anabaena sp
12	689.5	10.4	1197	2 Q7NJ67	Q7nj67 gloeobacter
13	677.5	10.2	1711	2 Q8Z019	Q8z019 anabaena sp
14	672	10.2	1683	1 Y124 ANASP	Q8yv57 anabaena sp
15	666	10.1	1193	2 Q7ND05	O7nd05 gloeobacter
16	659	10.0	1188	2 Q7ND80	O7nd80 gloeobacter
17	643.5	9.7	1356	1 HET1 PODAN	Q00808 podospora a
18	626	9.5	1693	1 Y163 SYN3	Q55563 synecocyst
19	624.5	9.4	1356	2 Q8X1P4	Q8x1p4 podospora a
20	614.5	9.3	934	2 Q8Y223	Q8yz23 anabaena sp
21	611.5	9.2	1551	2 Q8YMU3	Q8ymu3 anabaena sp
22	610	9.2	1708	2 Q8Y212	O8yi2 anabaena sp
23	609	9.2	1081	2 Q7NCT8	O7nct8 gloeobacter
24	606.5	9.2	1356	2 Q8X1P3	Q8x1p3 podospora a
25	605.5	9.1	1356	2 Q8X1P2	Q8x1p2 podospora a
26	604.5	9.1	1376	2 Q8X1P2	Q8x1p2 podospora a
27	600.5	9.1	1189	2 Q8Y109	O8yl09 anabaena sp
28	586	8.9	1183	2 Q7NLE9	O7nle9 gloeobacter
29	581	8.8	1233	2 Q8TMX4	Q8tmx4 methanosarc
30	580.5	8.8	1189	2 Q8YTD1	Q8ytd1 anabaena sp
31	561.5	8.5	1747	2 Q8Z020	Q8z020 anabaena sp

32	539	8.1	1241	2	Q9XBD8	Q9xbd8 amycolatops
33	531.5	8.0	1730	2	Q7NKS0	Q7nk50 gloeobacter
34	530.5	8.0	1193	2	Q7NH82	Q7nh82 gloeobacter
35	508	7.7	1051	2	Q8TMS3	Q8tms3 methanosarc
36	506	7.6	1671	2	Q7NMP0	Q7nmp0 gloeobacter
37	493.5	7.5	1721	2	Q7NID9	Q7nid9 gloeobacter
38	464.5	7.0	1329	2	Q8YYT7	Q8yyt7 anabaena sp
39	444	6.7	1538	2	Q6W219	Q6w219 rhizobium s
40	444	6.7	1538	2	AAQ87199	Aaq87199 rhizobium
41	442.5	6.7	1592	2	O62471	O62471 caenorhabdi
42	442.5	6.7	1592	2	CAAL6357	Caal6357 caenorhab
43	438	6.6	1237	2	Q6ZE54	Q6ze54 synecocyst
44	438	6.6	1237	2	BAD02046	Bad02046 synecocyst
45	423	6.4	2629	1	TEPI_RAT	O08653 rattus norv

ALIGNMENTS

RESULT 1

ID	APAF HUMAN	STANDARD;	PRT;	1248 AA.
AC	O14727; Q7Z438; Q9BXZ6; Q9UBZ5; Q9UGN8; Q9UGN9; Q9UGP0; Q9UJ58; Q9UJ59; Q9UJ60; Q9UJ61; Q9UJ62; Q9UJ63; Q9UJ64; Q9UJ65; Q9UJ66; Q9UJ67; Q9UNC9;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Apoptotic protease activating factor 1 (Apaf-1).			
GN	Name=APAF1; Synonyms=KIAA0413;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 3), AND PARTIAL SEQUENCE.			
RC	TISSUE=Cervical carcinoma;			
RX	MEDLINE=97410306; PubMed=9267021;			
RA	Zou H., Henzel W.J., Liu X., Lutschg A., Wang X.;			
RT	"Apaf-1, a human protein homologous to C. elegans CED-4, participates in cytochrome c-dependent activation of caspase-3.";			
RL	Cell 90:405-413(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1; 4 AND 5).			
RC	TISSUE=Cervical carcinoma, Heart, and Peripheral blood;			
RX	MEDLINE=99373149; PubMed=10441496; DOI=10.1006/bbrc.1999.1124;			
RA	Hahn C., Hirsch B., Janke D., Duerkop H., Stein H.;			
RT	"Three new types of Apaf-1 in mammalian cells.";			
RL	Biochem. Biophys. Res. Commun. 261:746-749(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=T-cell;			
RX	MEDLINE=99292765; PubMed=10364241;			
RA	Saleh A., Srinivasula S.M., Acharya S., Fishel R., Alnemri E.S.;			
RT	"Cytochrome c and darp-mediated oligomerization of Apaf-1 is a prerequisite for procaspase-9 activation.";			
RL	J. Biol. Chem. 274:17941-17945(1999).			
RN	[4]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, SUBUNITS, AND MUTAGENESIS OF			
RC	LYS-160 AND MET-368.			
RC	TISSUE=Kidney;			
RX	MEDLINE=99321747; PubMed=103931175;			
RA	Hu Y., Benedict M.A., Ding L., Nunez G.;			
RT	"Role of cytochrome c and darp/ATP hydrolysis in Apaf-1-mediated			
RL	caspase-9 activation and apoptosis.";			
RN	EMBO J. 18:3586-3595(1999).			
RN	[5]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RA	Ohara O.;			
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE OF 385-1248 FROM N.A. (ISOFORM 2).			
RC	TISSUE=Brain;			

KW	3D-structure; Alternative splicing; Apoptosis; ATP-binding; Direct protein sequencing; Repeat; WD repeat.	
KW	Query Match	
	Best Local Similarity 100.0%; Score 6619; DB 1; Length 1248;	
	Matches 1248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MDKARNCLQHREALDKDITSYTMDHMSDGLTISEEEKVRNEPTQOQRAAMLIKMI	60
Db	1 MDKARNCLQHREALDKDITSYTMDHMSDGLTISEEEKVRNEPTQOQRAAMLIKMI	60
Qy	61 LKQNDSSVYFYNALLHGYKDLAALLHDGIPWSSSGKDSVSGITSVRTVLCGGVP	120
Db	61 LKQNDSSVYFYNALLHGYKDLAALLHDGIPWSSSGKDSVSGITSVRTVLCGGVP	120
Qy	121 QRPVVVTRKLVNAIQKSLKGEKPGWVTIHGMAGCKSVLAEEAVRDSHLLSCFPFG	180
Db	121 QRPVVVTRKLVNAIQKSLKGEKPGWVTIHGMAGCKSVLAEEAVRDSHLLSCFPFG	180
Qy	181 GVHWVSQKQDGLLMLKQLNCTLDQDQESQRLPLNIEAKDRLILMLRKHPRSL	240
Db	181 GVHWVSQKQDGLLMLKQLNCTLDQDQESQRLPLNIEAKDRLILMLRKHPRSL	240
Qy	241 ILDDVWDSWVLKAFDSQCIILLTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFVN	300
Db	241 ILDDVWDSWVLKAFDSQCIILLTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFVN	300
Qy	301 MKKADLPQAHSIIECKGSPVLSLIGALLRDFPNRWEYVLQOLQNKQFKRIRKSSSYD	360
Db	301 MKKADLPQAHSIIECKGSPVLSLIGALLRDFPNRWEYVLQOLQNKQFKRIRKSSSYD	360
Qy	361 YEALDEAMSISVEMLRDIDKYITDLSILQKDVKPTKVLCLMDMETEEVEDILQEFVN	420
Db	361 YEALDEAMSISVEMLRDIDKYITDLSILQKDVKPTKVLCLMDMETEEVEDILQEFVN	420
Qy	421 KSLFLCDRNGKSFYRIYDLQVDFTEKNCQLODLHKKIITOFORYHQPHTLSPDQDC	480
Db	421 KSLFLCDRNGKSFYRIYDLQVDFTEKNCQLODLHKKIITOFORYHQPHTLSPDQDC	480
Qy	481 MYWYNFLAYHWSAKWHKELCALMFLSDWIKAKTELGVPAHLIHEFVYRHLDEKCAV	540
Db	481 MYWYNFLAYHWSAKWHKELCALMFLSDWIKAKTELGVPAHLIHEFVYRHLDEKCAV	540
Qy	541 SENFOEFLSLNHLGRLGPPFNIVOLGICEPETSSEVYQAKLQAKQEVNDGMLYLEWINK	600
Db	541 SENFOEFLSLNHLGRLGPPFNIVOLGICEPETSSEVYQAKLQAKQEVNDGMLYLEWINK	600
Qy	601 KNITNLSLVRPHPTDHYHACFSDEGQRIASCADKTLQVFKATGKLEIKAHEDV	660
Db	601 KNITNLSLVRPHPTDHYHACFSDEGQRIASCADKTLQVFKATGKLEIKAHEDV	660
Qy	661 LCCAFSTDDRFTATCSVDKVKIWNMTGELVHTYDEHSEOVNCCFTNSSHLLLATGS	720
Db	661 LCCAFSTDDRFTATCSVDKVKIWNMTGELVHTYDEHSEOVNCCFTNSSHLLLATGS	720
Qy	721 SDCFLKLWDLNQKCRNTMFGHTNSVNRHCRFSPPDKLLASCADGTLKLWATSANERKS	780
Db	721 SDCFLKLWDLNQKCRNTMFGHTNSVNRHCRFSPPDKLLASCADGTLKLWATSANERKS	780
Qy	781 INVQKFFLNLEDPOQDMVEIVKCCSWSDAGARIMWAAKNKIFLPIHTSGLLGEIHTGHH	840
Db	781 INVQKFFLNLEDPOQDMVEIVKCCSWSDAGARIMWAAKNKIFLPIHTSGLLGEIHTGHH	840
Qy	841 STIOYCFSPQNHLLAVVALSQYCVELWNTDSDRSKVADCRGHLVHWGVFSPDGSSFLTS	900
Db	841 STIOYCFSPQNHLLAVVALSQYCVELWNTDSDRSKVADCRGHLVHWGVFSPDGSSFLTS	900
Qy	901 SDDQIRLWETKVKCKSNVLMKQSDVVVFQNEVWMLAVDHIRLQNLINGRTGQIDYLT	960
Db	901 SDDQIRLWETKVKCKSNVLMKQSDVVVFQNEVWMLAVDHIRLQNLINGRTGQIDYLT	960
Qy	961 EAQVSCCLSPHLQVIAFGDENGAIIELELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS	1020
Db	961 EAQVSCCLSPHLQVIAFGDENGAIIELELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS	1020

RESULT 2

APAF_MOUSE			
ID	APAF_MOUSE	STANDARD;	PRT; 1249 AA.
AC	O88879;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Apoptotic protease activating factor 1 (Apaf-1).		
GN	Name=Apaf1;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RC	STRAIN=Swiss Webster/NIH Swiss; TISSUE=Embryo;		
RX	MEDLINE=98424242; PubMed=9753320;		
RA	Cecconi F., Alvarez-Bolado G., Meyer B.I., Roth K.A., Gruss P.;		
RT	"Apaf1 (CED-4 homolog) regulates programmed cell death in mammalian development.";		
RL	Cell 94:727-737(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RC	TISSUE=Spleen;		
RX	MEDLINE=21063441; PubMed=1119689; DOI=10.1016/S0006-8993(00)02916-4;		
RA	Walke D.W., Morgan J.I.;		
RT	"A comparison of the expression and properties of Apaf-1 and Apaf-1L.";		
RL	Brain Res. 886:73-81(2000).		
CC	- FUNCTION: Oligomeric Apaf-1 mediates the cytochrome c-dependent autocatalytic activation of pro-caspase 9 (Apaf-3), leading to the activation of caspase-3 and apoptosis. This activation requires ATP (By similarity).		
CC	- SUBUNIT: Monomer. Oligomerizes upon binding of cytochrome c and dATP. Oligomeric Apaf-1 and pro-caspase-9 bind to each other via their respective NH2-terminal CARD domains and consecutively mature caspase-9 is released from the complex (By similarity). It may also interact with Bcl-XL.		
CC	- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).		
CC	- ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=2;		
CC	Name=1; Synonyms=Apaf-1L;		
CC	Isoid=O88879-1; Sequence=Displayed;		
CC	Note=Major isoform;		
CC	Name=2;		
CC	Isoid=O88879-2; Sequence=VSP_006763;		
CC	- TISSUE SPECIFICITY: Highly expressed in lung and spleen, weakly in brain and kidney and not detectable in liver.		
CC	- DEVELOPMENTAL STAGES: High levels in embryonic brain and liver from E11.5 to E17.5 day.		
CC	- SIMILARITY: Contains 1 CARD domain.		
CC	- SIMILARITY: Contains 1 NB-ARC domain.		
CC	- SIMILARITY: Contains 13 WD repeats.		
CC	-----		

[illegible]

421	KSLLFCNENKGSFRYYLHDLOVDFTLETCNSQDLHKKIITQFORYHPHTLSPDOEDC	480
421	KSLLFCNENKGSFRYYLHDLOVDFTLETCNSQDLHKKIITQFORYHPHTLSPDOEDC	480
421	KSLLFCNENKGSFRYYLHDLOVDFTLETCNSQDLHKKIITQFORYHPHTLSPDOEDC	480
481	MTWYNFLAYHNASAKHKEI CALMESLDNI KAKTEL VGPANH IHEFVYRHLDEKDCAV	540
481	MTWYNFLAYHNASAKHKEI CALMESLDNI KAKTEL VGPANH IHEFVYRHLDEKDCAV	540
541	SNFOFFLSLNGHLLGRQFPFNI VOLGLCEPTESEVYQAKLOAQOEVDNGMLYEWINK	600
541	CENFOFFLSLNGHLLGRQFPFNI VGTGCEPTESEVYQAKLOAQOEVDNGMLYEWINK	600
601	KNITNLSLVVRPHTDAVYHACSDGQRIASCADKTLQVFKAPTGBKLEI IKAHDEV	660
601	KNITNLSLVVRPHTDAVYHACSDGQRIASCADKTLQVFKAPTGBKLEI IKAHDEV	660
661	LCCAFSTDRFIATCSYDKVKIWNMSITGELVHTYDHSHEQVNCCHFTNSSHILLATGS	720
661	LCCAFSSDYSIATCSADKKVKI WDSATGKLVHTYDHSHEQVNCCHFTNSSHILLATGS	720
721	SDCFPLKLDLNOKECRNTWFGHTNSVNHCRFSPDDKLLASCADGTGLKLWDATSANERKS	780
721	NDFFPLKLDLNOKECRNTWFGHTNSVNHCRFSPDDELLASCADGTGLRLWDVRSANERKS	780
781	INVKQFFNLNLEDOEDMEVIVKCCSWSADGARIWVAKNKI FLFDIHTSGLLGRHHTGH	840
781	INVKRFUSSEDPEDVEVIVKCCSWSADGDKI IVAKNKVLFDIHTSGLLAEIHTGH	840
841	STIQCDSPONHLAVVALSOYCVELWNTDRSKVADCRGHSWVGWVSPDGSSFLTS	900
841	STIQCDSPYDHLAVLALSOYCVELMNDLSRLKVADCRGHSWVGWVSPDGSSFLTA	900
901	SDDQIRLWETPKVCNKSAMVLKQEVVVUPOENEVWVLAVDHIRLQLINRGTCQIDYLT	960
901	SDDQIRVWETPKVCNKSAILVKQSDIVVPOENETWVLAVDNIIRGLQIAGTKTCQIDYLP	960
961	EAQVSCCLSPHQIYAFGBDNGAIEILLEVNNRI FQSRPHKTKVMHI QFTADEKTLIS	1020
961	EAQVSCCLSPHEIYAFGBDGAIKIIELPNNRVSSGVGHKCAVRHIQFTADGKTLIS	1020

[illegible]

RESULT 3			
AFAP_RAT	ID	AFAP_RAT	
AC	Q9BEF5	(Rel. 41, Created)	
DT	28-FEB-2003	(Rel. 41, Last sequence update)	
DT	28-FEB-2003	(Rel. 41, Last sequence update)	
DT	05-JUL-2004	(Rel. 44, Last annotation update)	
DT	28-FEB-2003		

DE Apoptotic protease activating factor 1 (Apaf-1).
GN Name=Apaf1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Itoh T., Itoh A., Pleasure D.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP DEVELOPMENTAL REGULATION, AND INDUCTION BY BRAIN INJURY.
RX MEDLINE=21450943; PubMed=11567033;
RA Yakovlev A.G., Ota K., Wang G., Movsesyan V., Bao W.-L., Yoshihara K.,
RA Faden A.I.;
RT "Differential expression of apoptotic protease-activating factor-1 and
RT caspase-3 genes and susceptibility to apoptosis during brain
RT development and after traumatic brain injury.";
RL J. Neurosci. 21:7439-7446(2001).
CC -!- FUNCTION: Regulates programmed cell death; necessary for normal
CC brain development. Participates with caspase-9 (Apaf-3) in the
CC cytochrome c-dependent activation of caspase-3, leading to
CC apoptosis. This activation requires ATP (By similarity).
CC -!- SUBUNIT: Monomer. Oligomerizes upon binding of cytochrome c and
CC dATP. Oligomeric Apaf-1 and pro-caspase-9 bind to each other via
CC their respective NH2-terminal CARD domains (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- DEVELOPMENTAL STAGE: Highly expressed in brain cortex in embryos
CC (E17) and new-born rats up to day 7. Very low expression
CC thereafter.
CC -!- INDUCTION: By brain injury.
CC -!- SIMILARITY: Contains 1 CARD domain.
CC -!- SIMILARITY: Contains 1 NB-ARC domain.
CC -!- SIMILARITY: Contains 13 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF320222; AAG35067.1; -.
DR HSP; P16649; 1ERJ.
DR RGD; 620575; Apaf1.
DR InterPro; IPR001315; CARD.
DR InterPro; IPR011029; DEATH_like.
DR InterPro; IPR002182; NB-ARC.
DR InterPro; IPR004106; Peptidase_S9A_N.
DR InterPro; IPR001680; WD40.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF00619; CARD; 1.
DR Pfam; PF00931; NB-ARC; 1.
DR Pfam; PF00400; WD40; 12.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 3.
DR SMART; SM00320; WD40; 13.
DR PROSITE; PS50209; CARD; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 4.
DR PROSITE; PS50082; WD_REPEATS_2; 9.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Apoptosis; ATP-binding; Repeat; WD repeat.
FT DOMAIN 1 90
FT NP_BIND 106 415 NB-ARC.
FT REPEAT 154 161 ATP (potential).
FT REPEAT 613 652 WD 1.
FT REPEAT 655 694 WD 2.
FT REPEAT 697 738 WD 3.
FT REPEAT 741 780 WD 4.
FT REPEAT 796 837 WD 5.
FT REPEAT 838 877 WD 6.

FT REPEAT 880 919 WD 7.
FT REPEAT 959 998 WD 8.
FT REPEAT 1001 1040 WD 9.
FT REPEAT 1042 1080 WD 10.
FT REPEAT 1083 1122 WD 11.
FT REPEAT 1125 1164 WD 12.
FT REPEAT 1176 1213 WD 13.
SQ SEQUENCE 1249 AA; 141151 MW; 7B4A8116FAD008E9 CRC64;
Query Match 88.7%; Score 5871.5; DB 1; Length 1249;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1091; Conservative 82; Mismatches 75; Indels 1; Gaps 1;
Qy 1 MDAKARNCLLOHREALEKDIKTSYIMDHIMISDGFITISEEKVNRNEPTQOQRAAMLIKMI 60
Db 1 MDAKARNCLLOHREALEKDIKTSYIMDHIMISDGFITISEEKVNRNEPTQOQRAAMLIKMI 60
Qy 61 LKKNDSSVSYFNALLHEGYKDLAALLHDGI PVVSSSSGKDSVSGITSVYRTVLCGGVP 120
Db 61 LKNONYAISYFNALLHEGYKDLAALLHSGGLPLVSSSSGKDTGGNTSFVRTVLCGGVP 120
Qy 121 QRPVVFVTRKKLVNAIQOKLSKLAGEPGWVTIHGMAGCGKSVLAEEAVRDHSLLEGCFPG 180
Db 121 QRPVVFVTRKKLVSAIQOKLWKLNGEPGWVTIYGMAGCGKSVLAEEAVRDHSLLEGCFPG 180
Qy 181 GVHWVSVKQDKSGILMKLQNLCTRLDQDESFSQRLPLNIEEAKDRLILMLRKHPRSL 240
Db 181 GVHWVSIGKQDKSGILMKLQNLCTRLGQESFSQRLPLNIEEAKDRLVLMRKHPRSL 240
Qy 241 ILDDVMSWLVKAFDSOCQILLTTTRDKSVTDSVNGPKYVVPVSSSLGKEGLEILSLFN 300
Db 241 ILDDVMDPWLKAFDNOCCILLTTTRDKSVTDSVNGPKYVVPVSSSLGKEGLEILSLFN 300
Qy 301 MKKADLPEQAHSIKECKGSLVVLGALRDPNRMWYLLKOLNQKFKRIRKSSSYD 360
Db 301 MKKEDLPVSAHSIKECKGSLVVLGALLRDPNRMWYLLRQNLQKFKRIRKSSSYD 360
Qy 361 YEALDEAMSIVEMLREDIKDYTTLSILOKDKVPTKVLCLMDMETEEVEDILQEPVN 420
Db 361 YEALDEAMSIVEMLREDIKDYTTLSILOKDKVPTKVLCLMDLETEEVEDILQEPVN 420
Qy 421 KSLIFCDNRNGSFYLLHLDQVDFLTNCSQLODLHKKIITQORVHOPHTLSPQSDC 480
Db 421 KSLIFCDNRNGSFYLLHLDQVDFLTNCSQLODLHKKIITQORVHOPHTLSPQSDC 480
Qy 481 MYWTFNLAYHMASAKMKELCALMFLSDWIKAETLVGPAHLIHEFVRYRHLDEKCAV 540
Db 481 MYWTFNLAYHMASAGMKELCALMFLSDWIKAETLVGPAHLIHEFVRYRHLDEKCAV 540
Qy 541 SENFQSLSLNGHLLGRQPPNIVOLGLCEPETSEVYQQAQKQAEVDNGLYLEWINK 600
Db 541 CENFQSLSLNGHLLGRQPPNIVOLGLCEPETSEVYQQAQKQAEVDNGLYLEWINK 600
Qy 601 KNIYNLSRLVVRPHTDAVYHACSEDQRIASCADKTLQVFKAEKGLEIKAEHDEV 660
Db 601 KTIKNLSRLVVRPHTDAVYHACSEDQRIASCADKTLQVFKAEKGLEIKAEHDEV 660
Qy 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
Db 661 LCCAFSSDDSYIATCSVDKVKIWNSTGKLVHTYEEHSEQVNCCHFTNSHHLLATGS 720
Qy 721 SDCFLKLWDLNQKCRNTMFGHTNSVNHCRFSPDDKLASCADGTLKLWDATSAERKS 780
Db 721 NDSFLKLWDLNQKCRNTMFGHTNSVTHCRFSPDDELLASCADGTLKLWDVRSANERKS 780
Qy 781 INVQKFLNLNLEDPQEDMEIVKCCSWADGARIWVAANKIFLPDIHTSGLLGEIHTGHH 840
Db 781 INVKRFLLSSDDPEDDEIVKCCSWADGRIIWAANKVLLDIHTSGLLTETHTGHH 840
Qy 841 STIQYCDSPONHLAVVALSQYVELWNDSRSKVADCRHLSVWGHVWFMFSPDGSSFLTS 900
Db 841 STIQYCDSPYDLHVALVALSQYVELWNDSRSKVADCRHLSVWGHVWFMFSPDGSSFLTA 900

```

QY 901 SDDQTRLWETKVKCNKSAVMLKQEVDDVFOENEMVAVLADHVRRLQRLNGRTQCIDYLT 960
DB 901 SDDQTRLWETKVKCNKSAVMLKQEVDDVFOENEMVAVLADHVRRLQRLNGRTQCIDYLT 960
QY 961 EAQVSCCLSPHLOVAFDENGAEIELELVNNRIFQSRFOHKTVMHIOFTADEKTLIS 1020
DB 961 EAQVSCCLSPHLEVAFDENGAEIKIIELPNNRVFSSGIGHKAVRHIOFTADGKTLIS 1020
QY 1021 SSDDAEIQVNNWQDKCIFLRGHOETVKDFRLLKNSRLLSWSFDGTVKVMNIIITGNKEKD 1080
DB 1021 SSDDAEIQVNNWQDKCIFLRGHOETVKDFRLLKNSRLLSWSFDGTVKVMNIIITGNKEKD 1080
QY 1081 FVCHQGVLSCDISHDAYTKFSTGADTKAKIWSFDLPLHRLRHGNCVRCSAFSDVST 1140
DB 1081 FVCHQGVLSCDISHDAYTKFSTGADTKAKIWSFDLPLHRLRHGNCVRCSAFSDVST 1140
QY 1141 LLATGDNGEIRIWNVSGELLHLHCAPLS-EEGAATHGGWVTDLCFSPDGKMLISAGGYI 1199
DB 1141 LLATGDNGEIRIWNVSGELLHLHCAPLS-EEGAATHGGWVTDLCFSPDGKMLISAGGYI 1199
QY 1200 KWNVNVYGRSSQTEVYNGTNIKKTHVSPDKTVVTDNLGILYILOTLE 1248
DB 1201 KWNVNVYGRSSQTEVYNGTNIKKTHVSPDKTVVTDNLGILYILOTLE 1249

RESULT 4
QSVI66 PRELIMINARY; PRT; 1249 AA.
AC QSVI66; 2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Apoptotic protease activating factor-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum.
RA Cao G., Chen J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC EMBL; AF218388; AAL36935.1; -.
DR HSP; P16649; IERJ.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0002333; F:peptidase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0043829; P:defense response to pathogen; IEA.
DR GO; GO:0042981; P:regulation of apoptosis; IEA.
DR InterPro; IPR001315; CARD.
DR InterPro; IPR011029; DEATH_like.
DR InterPro; IPR000767; Disease_resist.
DR InterPro; IPR002182; NB-ARC.
DR InterPro; IPR001680; WD40.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF00619; CARD; 1.
DR Pfam; PF00931; NB-ARC; 1.
DR Pfam; PF00400; WD40; 12.
DR PRINTS; PR00364; DISEASERST.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 3.
DR SMART; SM00320; WD40; 13.
DR PROSITE; PS02029; CARD; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 4.
DR PROSITE; PS00082; WD_REPEATS_2; 9.
DR PROSITE; PS02094; WD_REPEAT_REGION; 1.
KW Protease; Repeat; WD_repeat.
SQ SEQUENCE 1249 AA; 141169 NW; 2B41B629599A9A38 CRC64;

```

Query Match 88.6%; Score 5863.5; DB 2; Length 1249;
 Best Local Similarity 87.2%; Pred. No. 0;

```

Matches 1089; Conservative 84; Mismatches 75; Indels 1; Gaps 1;
QY 1 MDAKARNCLLOHREALEFKDITKYSIMDMHISDGFITSEEEKVRENEPTQOQAAAMIKMI 60
DB 1 MDAKARNCLLOHREALEFKDITKYSIMDMHISDGFITSEEEKVRENEPTQOQAAAMIKMI 60
QY 61 LKNDNSVSVSYNALLHEGYKDLAALLHDGIPVSSSSGKDSVSGITSYVYRTVLCGGVP 120
DB 61 LKNDNSVSVSYNALLHEGYKDLAALLHDGIPVSSSSGKDSVSGITSYVYRTVLCGGVP 120
QY 121 QRPVVFTRKKLVNAIOOKLSLKGEPPGWTIHGMAGCGKSVLAABAVRDHSLLECGPFG 180
DB 121 QRPVVFTRKKLVNAIOOKLSLKGEPPGWTIHGMAGCGKSVLAABAVRDHSLLECGPFG 180
QY 181 GVHWVSKQDKSGLLWKLNQCTRLDODESPORLPLNTEEKADRLILMLRKHPRL 240
DB 181 GVHWVSKQDKSGLLWKLNQCTRLDODESPORLPLNTEEKADRLILMLRKHPRL 240
QY 241 ILDDVWDSWVLKAPDSQCQILLTTRDKSVTSVMGPKYVVPVSSSGKEKGLIILSLFN 300
DB 241 ILDDVWDSWVLKAPDSQCQILLTTRDKSVTSVMGPKYVVPVSSSGKEKGLIILSLFN 300
QY 301 MKKADLPQAHSIIECKGSPVWSLIGALLRDPNNWEYLLKOLQNKOPKRIKSSSYD 360
DB 301 MKKADLPQAHSIIECKGSPVWSLIGALLRDPNNWEYLLKOLQNKOPKRIKSSSYD 360
QY 361 YEALDEAMISVEMLRDINKDYITDLSILOKDVKPTKVLCLWDMETEVEDILOEPVN 420
DB 361 YEALDEAMISVEMLRDINKDYITDLSILOKDVKPTKVLCLWDMETEVEDILOEPVN 420
QY 421 KSLFPCDRNGKSPRYLHDLQVDFLTEKNCSQLDOLHKKIITOFYHOPHTLSPDQDC 480
DB 421 KSLFPCDRNGKSPRYLHDLQVDFLTEKNCSQLDOLHKKIITOFYHOPHTLSPDQDC 480
QY 481 MYWNYFLAYHMAKQHKELCALMFSLDWIKAKTELGVPAHLHIEFVYVRHILDEKCAV 540
DB 481 MYWNYFLAYHMAKQHKELCALMFSLDWIKAKTELGVPAHLHIEFVYVRHILDEKCAV 540
QY 541 SENFOEFLSNGHLGRQPPNIVOLGCEPETSEVYQAKLOAKOEVDNGMYLEWINK 600
DB 541 SENFOEFLSNGHLGRQPPNIVOLGCEPETSEVYQAKLOAKOEVDNGMYLEWINK 600
QY 601 KNITNLSRLVVRPHTDANVHACFSDGRIASCADTKTLOVFAETGEXLLEIKAHEDV 660
DB 601 KNITNLSRLVVRPHTDANVHACFSDGRIASCADTKTLOVFAETGEXLLEIKAHEDV 660
QY 661 LCCAFSTDRTATCSVDKVKIKNSTMTELAVHTYDEHSEVNCCHFTNSSHILLATGS 720
DB 661 LCCAFSTDRTATCSVDKVKIKNSTMTELAVHTYDEHSEVNCCHFTNSSHILLATGS 720
QY 721 SDCFLKLDLWLNQKECNTWFGHTNSVNHCRFSPDDKLLASCADGTLLWDATSANERKS 780
DB 721 SDCFLKLDLWLNQKECNTWFGHTNSVNHCRFSPDDKLLASCADGTLLWDATSANERKS 780
QY 781 INVKOFFLANLEDPOEDMEVIVKCCSWSADGARIMVAARKNIFLFDIHTSGLLGEIHTGHH 840
DB 781 INVKOFFLANLEDPOEDMEVIVKCCSWSADGARIMVAARKNIFLFDIHTSGLLGEIHTGHH 840
QY 841 STIQYCFSPQNHAVVALSQYCVELMNTDSRKVADCRGHLSSWHGVNMFSPDGSFLTS 900
DB 841 STIQYCFSPQNHAVVALSQYCVELMNTDSRKVADCRGHLSSWHGVNMFSPDGSFLTS 900
QY 901 SDDQTRLWETKVKCNKSAVMLKQEVDDVFOENEMVAVLADHVRRLQRLNGRTQCIDYLT 960
DB 901 SDDQTRLWETKVKCNKSAVMLKQEVDDVFOENEMVAVLADHVRRLQRLNGRTQCIDYLT 960
QY 961 EAQVSCCLSPHLOVAFDENGAEIELELVNNRIFQSRFOHKTVMHIOFTADEKTLIS 1020
DB 961 EAQVSCCLSPHLEVAFDENGAEIKIIELPNNRVFSSGIGHKAVRHIOFTADGKTLIS 1020
QY 1021 SSDDAEIQVNNWQDKCIFLRGHOETVKDFRLLKNSRLLSWSFDGTVKVMNIIITGNKEKD 1080
DB 1021 SSDDAEIQVNNWQDKCIFLRGHOETVKDFRLLKNSRLLSWSFDGTVKVMNIIITGNKEKD 1080

```


Qy 1081 FVCHQTVLSCDISHDATKFSSTGADTKATKWSFDLPLPLHRLGHNGCVRCSAFSDVST 1140
 Db 1081 FTCHQTVLSCAISDATKFSSTGADTKATKWSFELSPHLGHNSCVRCSAFSLDGI 1140
 Qy 1141 LLATGDDNGEIRIWNVSGELLHLICAPUS-EGGAATHGCGWVTDLCFSPDGKWLISAGGYI 1199
 Db 1141 LLATGDDNGEIRIWNVSDQGLHLICAPISIEGATHGCGWVTDVCFSPDRKWLVSAGYL 1200
 Qy 1200 KWNVVTGESSQTFTYNTGNTLKKIHVSPDFKTYTYVDNLGLYILQTL 1248
 Db 1201 KWNVVTGESSQTFTYNTGNTLKKIHVSPDFRTYTYVDNLGLYILQVLE 1249

RESULT 5
 Q6GNU6 PRELIMINARY; PRT; 1248 AA.
 AC Q6GNU6;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heideh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 initiative."
 RL Dev. Dyn. 225:384-391(2002).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S., Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 11 WD repeats.
 DR EMBL; BC073405; AAH73405.1; -.
 DR InterPro; IPR01315; CARD.
 DR InterPro; IPR011029; DEATH like.
 DR InterPro; IPR000767; Disease resist.
 DR InterPro; IPR002182; NB-ARC.
 DR InterPro; IPR001680; WD40.
 DR InterPro; IPR011046; WD40 like.

DR Pfam; PF00619; CARD; 1.
 DR Pfam; PF00931; NB-ARC; 1.
 DR Pfam; PF00400; WD40; 11.
 DR PRINTS; PR00364; DISEASERSIST.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR SMART; SM000018; WD40; 2.
 DR SMART; SM00320; WD40; 13.
 DR PROSITE; PS02029; CARD; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 3.
 DR PROSITE; PS00882; WD_REPEATS_2; 9.
 DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
 KW Hypothetical protein; Repeat; WD repeat.
 SQ SEQUENCE 1248 AA; 141574 MW; 822F7CEAC4EE675D CRC64;

Query Match 61.2%; Score 4054; DB 2; Length 1248;
 Best Local Similarity 60.5%; Pred. No. 3e-227;
 Matches 757; Conservative 198; Mismatches 289; Indels 8; Gaps 4;

Qy 1 MDKARNCLLOHREALEKDIKTSYIMDHMSDGLFTISESEKVRNEPTQOORAAWLKMI 60
 Db 1 MDEKARSLLLQNRALTALVRDITAYIMDHMSDQVITPEEARVKAQHTQVDKANFLNLI 60
 Qy 61 LKQNDSDYVSFYNALLHEGYKDLAALLHDGIPVVSSSSGKSDSVSGITSYVTVLCEGVP 120
 Db 61 LGKGQAYVSFYNALLHEGYKDLALLKEAANAHSSTKSSNGITSYVTVLCEGAVP 120
 Qy 121 QRPVVFVTRKLVNAIQOKLSKLGEFGWVTIHMAGCGKSVLAEEAVRDHSLLEGCFPG 180
 Db 121 QRPVVFVTRPKLVTKIQOSLYKLVNVEGWTVTYMGAGCGKSVLAEEAVRDHSHKILTECFPG 180
 Qy 181 GVHWVSVGKODKSGLLMKLQNLCTRLDQDSFSORLPINIEEAKDRILMLRKHPSLL 240
 Db 181 GVHWVSVGKODKSGLLMKLQNLCTRLDOEAKYSQRPFLNAEAEARERLMLVMVNRPRCLI 240
 Qy 241 ILDDVDWSWLKAFDSOCOLLTTTRDKSVTDSVVGPKVVPVSVSSSLGKEGLEILSLFVN 300
 Db 241 VLDDVDWSWLKAFDPIQORVLIITRDKSVTDSLSGFEAVRVDSLEHSGKEGLEILSLFVS 300
 Qy 301 MKKADLPEQAHSIKECKGSPVLSLIGALLRDPNRPWEYLYLKQLONKQFKRIRKSSSYD 360
 Db 301 LKABELPSQAHSIKVKSGSPVLSLIGALLRDEPFRWDFYLTQKRRKQPKRIRKSSSYD 360
 Qy 361 YEALDEAMSSIVEMLRDIDKYDTSILQKDVKVPKVLILWDMTEVEVEDILOEFPVN 420
 Db 361 YEALDEAMSSIVDNKLEDFREYNDFSIIEKDVKLPTQVLCILWDMEREDVEDMMQEPVN 420
 Qy 421 KSLILFCDRNGKSPRYLHLDQVDFLTEKNCSQLQDLHKIITQFQYHQPHTLSPDOEDC 480
 Db 421 KSLIYCDRNGKSPRYLHLDQVDFLTEKNCSQLQDLHLDQVDFLTERNRDLTVLHAKLVQYKXHYSSKLPVEQEDC 480
 Qy 481 MYWYNFLAYHMASAKMKEICALMFSLDWIKAETELVGPAPHLIHEFVEYRHLDEKCAV 540
 Db 481 VYWHYLAHYMAKANLHQLCSLLFSLNLWLNKAKSELFTGSHLHIEFVQYRNILQNAKA 540
 Qy 541 SENPOEFLSINGHLHGQPPFNIVQLGCEPTESEVYQOAKLOAKQOEVDNGMLYLEWINK 600
 Db 541 REHFQEPYAVNGHLLGNLPDISIVQLGLGAEDSEVYKQAILQAKELKSGVLVYNWINK 600
 Qy 601 KNITNLRLVVRPHDTAVYHACFSEDCGRIASCADKTLQVFKAEKGKLEIKAHEDV 660
 Db 601 QRMKNLPTLVVRPHKDAVHACFSPDGRKFPASCADKTLQVFRSETEKLELEAHDDV 660
 Qy 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDSEHSEQVNCCHFTNSHHLLLATGS 720
 Db 661 LCCAFSADEKLLATCSADRKVKIWNKTKPIRVYEEHTEQVNCCTFNGLSAPLLATCS 720
 Qy 721 SDCLPKLWLNQKCEKRTMFGHTNSVNHCRFSPDDKLLASCSDGTLKLWDATSAERKS 780
 Db 721 NDCPILMLWDSSETYSRNTLIGTCAVYHCRYSPPDRYLASSMDGSLKIWDVESANBEKS 780
 Qy 781 INVKQFPLNLEDPQEDMEVIVKCSWSADGARTMVAANKIFLFIHTSGILGHIHTGH 840
 Db 781 IEVAKLFENEDESQF--EVLLKCCAWNSNDGSRIMVTRNLFICIFDSTSCDLLSLQKACHQ 838

QY 841 STIQYCDSPQNHLLAVVALSQYCVELWNTDSRSKVADCRGHLSSWHVGHMSPDGSSFLTS 900
Db 839 --ILYCDFCTTQNIQVALALSYYVQLWMDISDTKIAFNAHLSWVHCYKSPKSSFLTS 896
QY 901 SDQTLRLWETKVKCKSAVNLKQEVVVPOENEVNVLAADVHRRLOQLINGRGQIDYLT 960
Db 897 SDDQTVKLWETSNVSKPSATNLKREFDVSPNGEETLVLATSKDDCILLINGMTGETLSQI 956
QY 961 EAG--VSCCLSPHLYOIAFGDENGAGIELELVNRRIFQSFQHKTKVWHIQTADKX 1017
Db 957 NTQDKCVTCCLTNDYQLAAGDREDGKVIDVSRGEILCKLGDHSATVQCFTADGKH 1016
QY 1018 L1SSSDARIQVWVWQDKCIFRCHOETVKQFRLLKNSRLLSNFDGTVKWNITGNK 1077
Db 1017 LVSSSDSTIRVWSLASGESLELHGKPEVPFKLQSQVFSFDTGVKVDLITGKL 1076
QY 1078 EKDFVCHQGTULSCDISHDATKSPSTADTKAKINSFDFLLPLHELCHNGCVRCSAFSV 1137
Db 1077 RKEFICHSETVLSCDISPDSTKFSADKSKINSLDMSTLLHGLHGHQSCVRCRFSW 1136
QY 1138 DSTLLATGDNGBEIRIMVNSVNGELLHLCAPLSEBGA-ATHQGWVTDLCFPGDGRMLISAG 1196
Db 1137 DNKYLATGDNGKIMNSVQNGELLKQCCDISVNNENSLHGDGWVTDLHFSNPKLIVSSG 1196
QY 1197 GYIKWNVVTESSQTFYNTGNLKKIHVSPDKTYVVDNLGILYILOTLE 1248
Db 1197 ANYKXWVDVDTGACKQTFYVTGNLKSILHLCPTFTFTVTDNLGILYVCLKLIE 1248
RESULT 6
APAF BRARE STANDARD; PRT; 1261 AA.
AC Q919H8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apoptotic protease activating factor 1 (Apaf-1).
GN Name=apaf1
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2037392; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
R1 Zebrafish.";
R1 Cell Death Differ. 7:509-510(2000).
CC -!- FUNCTION: Oligomeric Apaf-1 mediates the cytochrome c-dependent
CC autocatalytic activation of pro-caspase 9 (Apaf-3), leading to the
CC activation of caspase-3 and apoptosis. This activation requires
CC ATP (By similarity).
CC -!- SUBUNIT: Monomer. Oligomerizes upon binding of cytochrome c and
CC dATP (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 CARD domain.
CC -!- SIMILARITY: Contains 1 NB-ARC domain.
CC -!- SIMILARITY: Contains 13 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF251502; RAJ67189.1; -.
DR HSPSP; P16649; 1ERU.
DR ZFIN; ZDB-GENE-0000616-4; apaf1.

InterPro; IPR001315; CARD.
DR InterPro; IPR00182; NB-ARC.
DR InterPro; IPR001680; WD40.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF00619; CARD; 1.
DR Pfam; PF00931; NB-ARC; 1.
DR Pfam; PF00400; WD40; 11.
DR PRINTS; PR00320; GPROTEINERPT.
DR ProDom; PD000018; WD40; 2.
DR SMART; SM00114; CARD; 1.
DR SMART; SM00320; WD40; 14.
DR PROSITE; PSS0209; CARD; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PSS0082; WD_REPEATS_2; 7.
DR PROSITE; PSS0294; WD_REPEATS_REGION; 1.
KW Apoptosis; ATP-binding; Repeat; WD repeat.
FT DOMAIN 1 90
FT NP_BIND 106 417 NB-ARC.
FT REPEAT 154 161 ATP (potential).
FT REPEAT 615 654 WD 1.
FT REPEAT 657 696 WD 2.
FT REPEAT 700 743 WD 3.
FT REPEAT 746 785 WD 4.
FT REPEAT 798 836 WD 5.
FT REPEAT 840 879 WD 6.
FT REPEAT 882 921 WD 7.
FT REPEAT 964 1003 WD 8.
FT REPEAT 1006 1045 WD 9.
FT REPEAT 1047 1088 WD 10.
FT REPEAT 1091 1130 WD 11.
FT REPEAT 1133 1172 WD 12.
FT REPEAT 1184 1223 WD 13.
SQ SEQUENCE 1261 AA; 142933 MW; F80CC9CB41764F1C CRC64;
Query Match 55.7%; Score 3683.5; DB 1; Length 1261;
Best Local Similarity 54.8%; Pred No. 1.2e-205;
Matches 691; Conservative 244; Mismatches 310; Indels 15; Gaps 7;
QY 1 WDARVNCILQHREALEKDKTSYIMDMHISDGLTITSEEEKVNEPTQOORAMLIKWI 60
Db 1 MEERARELLRSKATLEQDIKASYLMDHISDGLTITSEEEKVNEPTQOORAMLIKWI 60
QY 61 LKQNDSTVYFNALLHEGYKDLAALLHGDGIPVSSSGKDSVSGTISYRTVLCGGVP 120
Db 61 LRKONRAYISFYNALITRESYGDLASLLHSDLPSPGKSFADGVSVPVQALLSGVGP 120
QY 121 QRPVVVTRKLVNVAIOOKLSKLGEPGWTHIGHMAGCGKSVLAABAVRDHSLLEGCPFG 180
Db 121 QRPVVVTRKLVNVAIOOKLSKLGEPGWTHIGHMAGCGKSVLAABAVRDHSLLEGCPFG 180
QY 181 GVHVSVSGKQKSGLLMLQLNLCRLDQDES--FSQRLPLNIEEAKDRILRLMRKPR 238
Db 181 GVHVSVSGKQKSGLLMLQLNLCRLDQDES--FSQRLPLNIEEAKDRILRLMRKPR 238
QY 239 LLILDDVWDSWVLKAFDSQCQILLTRDKSVTDSVMGPKVVPVSSSLGKGLTSLF 298
Db 241 LLILDDVWDSWVLKAFDSQCQILLTRDKSVTDSVMGPKVVPVSSSLGKGLTSLF 298
QY 299 VMKKADLPQAHSIIECKGSPVLSLIGALLRDFPNRWEYVYLKOLNOKFKRIRKSS 358
Db 301 VNGRMKLPQAHSIIECKGSPVLSLIGALLRDFPNRWEYVYLKOLNOKFKRIRKSS 358
QY 359 VYDALDEMSISVEMLRDIDKDYVTDLSILQDKVPTKVLCLMDMETEVEDILQEF 418
Db 361 VYDALDEMSISVEMLRDIDKDYVTDLSILQDKVPTKVLCLMDMETEVEDILQEF 418
QY 419 VNKSLICDNGKSPFYVLDHLDVDFLEKSGVLOLHKKLTQORVHOPHTLSPOE 478
Db 421 VNKSLICDNGKSPFYVLDHLDVDFLEKSGVLOLHKKLTQORVHOPHTLSPOE 478
QY 479 DCMVTYFLAYHASAKMELCALMFLSDWKATKTELVPALHIEFVEYRHLDEKDC 538
Db 481 DCMVTYFLAYHASAKMELCALMFLSDWKATKTELVPALHIEFVEYRHLDEKDC 538
QY 540 DLSYVYQFIPYHMAKAGLSKELYSLMFSLDWKVKARIMGSAHLINDYVEYGEILDKENS 540

QY 539 AVSENFQFSLNGHLLGRQPPNIVQLGCEPSETSEVYQAAQKQAEVNGMYLEWI 598
 DB 541 EVRVQFQSLNGHLLGRQPPNIVQLGCEPSETSEVYQAAQKQAEVNGMYLEWI 600
 QY 599 NKNITNLRLVVRPHPTDAVYHACSEDPORIASCGADKTIQVPAKGEKLEIKAHED 658
 DB 601 KNTEELGSLRVLVMPHQAVYACFKDSKIASCGASKALURVFASTSGEKLELQAHEE 660
 QY 659 EVLCAFSDFRFTATCSVDKVKVWISMTGELVHYD-EHSEQVNCCHFTNSHLLLA 717
 DB 661 DVLCAFSDFRHTATCASDRKVLNVRGVLRFEFEVEHEEQINHCQFTNGRRVLLA 720
 QY 718 TGSSDCP--LKLWLNQKECRNTWPGHTNSVNHCRFPSPDDKLLASCADGTILKWDATSA 775
 DB 721 TCSNDKFTNTLWPNKKTQNTWPGHMEPVNHCFCSPNDLYLATSSDGLKLEFVSSA 780
 QY 776 NERKSNVQFPLNLEDPQEDMEVIVKCCSADGARIWAAKNIPLFHTHTSGLGEI 835
 DB 781 NEWKSLDVSFPF---PESDEBIKAMVKCSTWSDGSIICARNTVFVDFVETSDLLKL 837
 QY 836 HTGHSTIQYCDQFQNHLLAVVLSQYCVELWNTDSRSKVDRCRGLSHVWGVFSPDGS 895
 DB 838 KTSRLSTIQFCHACPNSSLLAVLSHYTVELWNPESKKKAECSGHLVWVHCVFSPDGS 897
 QY 896 SFLTSSDDQTLRLWETKVKCNKSAVMLKQEVVDVFOENEVMVLAVDHIRRLQLNGRTGQ 955
 DB 898 LLLSSDDQTLRLWETDRVHTSSAVALKRDVLLSSHSDATIIAPDSSNRLQLVSGSTGA 957
 QY 956 IDYLTE---AQVSCCLSPHLQYAFDENGAEIIELELVNRIQSRFQHKKTWHLOFT 1012
 DB 958 VVLESEELSSRCSIRSNAAFGALGSEDGTQVIEVPSSKASVKLSGGHTKTVHHCOFT 1017
 QY 1013 ADEKTLISSDDAEIQVNWOLDKICIFLRHQETVKDPFLKNS---RLLSWSPDGTVKV 1069
 DB 1018 DDCILLITSSDSTIRVWKTGECVWLGHEPVRKEPHLLSSSSPHLFSWSPDGTVKV 1077
 QY 1070 WNIITGNKEKDFVCHQGTVLSGDIASHDATKFSSTSDAKTAKIWSFDLLPLHLHLRGNCG 1129
 DB 1078 WDLTRGQMLQDLVCEGAVLSCDVSDDGRLEFATTSAKTAQVWSSASWMLFLEGHKDC 1137
 QY 1130 VRCSAFSDVSTLLATGDDNGIRIWNVNGELLHLCAPLSEGA-ATHGGVNTDLCPSPD 1188
 DB 1138 VRSCRFSDNRLATGDDNGIRIWSMDGALLKICPRDTKDSMNSVHAGVNTDLHSPD 1197
 QY 1189 GKMLISAGYIKWNVVTVGESSQFTYGTGTLKTHVSPDKTYVTVDNLGLIYLOLE 1248
 DB 1198 NRVLVSTAGYIKWNVVTVGESSQFTYGTGTLKTHVSPDKTYVTVDNLGLIYLOLE 1257

RESULT 7
 Q80VR5 PRELIMINARY; PRT; 258 AA.
 AC Q80VR5
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Similar to apoptotic protease activating factor 1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RN RN
 RC TISSUE=Olfactory epithelium;
 RX MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore S.I., Max A.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalls D.B., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RC SEQUENCE FROM N.A.
 RA TISSUE=Olfactory epithelium;
 RA Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC045144; AAH45144.1; -
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0005234; F:ATP binding; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.
 DR InterPro; IPR001315; CARD.
 DR InterPro; IPR011029; DEATH like.
 DR InterPro; IPR002182; NB-ARC.
 DR Pfam; PF00619; CARD; 1.
 DR Pfam; PF00931; NB-ARC; 1.
 DR PROSITE; PS02029; CARD; 1.
 KW Protease.
 SQ SEQUENCE 258 AA; 28808 MW; E1C16DEAD1D370A9 CRC64;
 Query Match 16.1%; Score 1066; DB 2; Length 258;
 Best Local Similarity 86.1%; Pred. No. 3.6e-54;
 Matches 204; Conservative 18; Mismatches 15; Indels 0; Gaps 0;
 QY 1 MDAKARNCLLQHREALEKDKITSYIMDMHISDGLTITSEEEKVNEPTQQQRAAMLKMI 60
 DB 1 MDAKARNCLLQHREALEKDKITSYIMDMHISDGLTITSEEEKVNEPTQQQRAAMLKMI 60
 QY 61 LKQNDSDVSVFYNNALHEGKDLAALLHDGTPVYSSSSGKDSVSGITSVTVLCEGVP 120
 DB 61 LKQNDSDVSVFYNNALHEGKDLAALLHDGTPVYSSSSGKDSVSGITSVTVLCEGVP 120
 QY 121 QRPVVFVTRKLVNAIQOKLSKLGEPGWTVIHGMAGCGKSVLAEEAVRDSHSLGCGPFG 180
 DB 121 QRPVVFVTRKLVNAIQOKLSKLGEPGWTVIHGMAGCGKSVLAEEAVRDSHSLGCGPFG 180
 QY 121 QRPVVFVTRKLVNAIQOKLSKLGEPGWTVIHGMAGCGKSVLAEEAVRDSHSLGCGPFG 180
 DB 121 QRPVVFVTRKLVNAIQOKLSKLGEPGWTVIHGMAGCGKSVLAEEAVRDSHSLGCGPFG 180
 QY 181 GVHWVSVGKQKSGLLMKLQNLCTRLDQDSFSQRLPLNTEBEAKDRILMLRKHPR 237
 DB 181 GVHWVSVGKQKSGLLMKLQNLCTRLDQDSFSQRLPLNTEBEAKDRILMLRKHPR 237

RESULT 8
 Q820R1 PRELIMINARY; PRT; 1227 AA.
 ID Q820R1
 AC Q820R1
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE WD-40 repeat protein.
 GN OrderedLocusNames=alr0029;
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 NCBI_TaxID=103690;
 [1]
 RN RN
 RC SEQUENCE FROM N.A.
 RX MEDLINE=2159285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,

Query Match 10.2%; Score 677.5; DB 2; Length 1711;
Best Local Similarity 22.8%; Pred. NO. 2.5e-30;
Matches 322; Conservative 248; Mismatches 527; Indels 315; Gaps 64;

QY 7 NCLQHREALEKDKTSYIMDHMTSDGFLTISEBEKVRNEPT-----QQQ 51
DB 359 NC-LQSR-LTENQONLATLL--YLRKLAESPTEGPVLVGENIQLNRWLROOL 413

QY 52 RAAMLIRKML-----KKONDSYVSFYNALHGYKDLAALLHDGIPVSSSGKDSVSGIT 107
DB 414 RRSLSYQIIFDCPLDKHSYIS-----LDWVEDLQGFQGCIIIAAASPENPQOFL 468

QY 108 SYRTVL-----CEGVPQRPVVFVTRKLVNA 135
DB 469 QTLHSTLQAHQEPNLSAAWINQLSPPLHLWLSGARGVIEIIPASTDAKGRPNA 528

QY 136 I-----CQKSLKLGEPGWTHGMAGGKSV--- 162
DB 529 IVDLGI CYPYRGLQAFQEDVQYFGRETLTQOLADLANK-SFMAVVGASGGKSVVQA 587

QY 163 -LAEAARVDSLELGGCFPGVHVSVGKQDKSGLLMKLN-LCTRLDQDESFSQRLPLN- 219
DB 588 GLIAQLRGGQL-----PGSEWMMKSPRPGEPYPLVSLCHLVDSGTEKEKAYQOMLEG 642

QY 220 --IEBAKRLRLMLRKHPRSLILLDDVMDSWLKAQDSQCQIILT-----TRDK--- 267
DB 643 MLYOGAOGFVHWHHRREPMVVLVDQPEELFTLAASEDRQRFDTVLGALELSPDKPKL 702

QY 268 --SVTDSVMGPKYVVPVSSIGKEGL-----ETLSLFVN-MKKADL---PEQAHS 312
DB 703 IVTLRAPIACLEIPLTAKLQOSSILLPCLTQOEYRRIIHPAEKVLGTVDPELVEV 762

QY 313 IIEKCKSPVLSVIGALLRDPFNRWEY-----YKQLQKQKPKRKRKSSSY 359
DB 763 LLOELHNSPGDLPLEFVLEQL---WEYRDKGVITLQAYQYLGIGK---ALEKKAQG 815

QY 360 DYEALD-EAMS-----ISVEMLRD-----IKDYTLDSILQKDVKVPYTK 398
DB 816 VYDTLDSEAQCTRWIFLSLTQLGEGTETRRRLVLSKELIVKKY--PVALVERTLQVLT 873

QY 399 VLCILMDMETEE-----VEDILQ-----EFVYKSLFCDRNGKSFYYLH 438
DB 874 AKLVVNGEWEEGKQKQAGGROGGENILTTSPVTEVAHEVII---RYWSTLRWL- 929

QY 439 DLQVDFLTEKNCSQLQDLHKIITQFORYPHPTLSPD-----QEDCM 481
DB 930 -----EENRSLRS-HRQ-IEQSAALWQNNQPDFFLQGVRLAEAEIYLYNDEL 979

QY 482 YW--YNFL--AYHWASAKMHKELCALMPSLDWIKAKTEL-----VGPALHIEFVYRHIL 533
DB 980 SWDVQHFTEACLHERRRQKQEQSRQAQRAVSIISTGLTAFGLAVFAVYQQTQNAQLK 1039

QY 534 DEKDC-AVSENP-----QEFSLNGLHLCRQFPFNVOLGL---CEPETSEYVQAKLQAK 585
DB 1040 BIQALNSFLNLSHQLEALMTSVQAGKEVQNI-SLIGPPDTRTQTATLQOAVNTQ 1098

QY 586 QEVNDGMLYLEWINKNITNLSRLVVRPHPTDAVYHACFSEGOIRASCGADKTLQVFAE 645
DB 1099 ER--NRLLHNAW-----VTSVS-----YSPGEVIVASGVNDTHLWRD 1136

QY 646 TGEKLRIKAHEDVLCCAFSTDTRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNC 705
DB 1137 -GKLLTTLTGNDGVNSVSPDCEILLASASADSTIKLW-QRNGQLITLKGHDQGVKSV 1194

QY 706 HFTNSSHLLLATGSSDCFLKWLNDNQECRNTWFGHTNSVNHCRFPSPDDKLLASCADG 765
DB 1195 SPSNGE--IIASQSSDHTINLWSRAGK-LLLSLNGHSQGVNSIKFSPGDTIASASDDG 1251

QY 766 TLKLWATSANERKSINVQKFFLNLDPQEDMEVIVKCCSWASDAGTMA-AKKKIFLF 824
DB 1252 TIRLW--SLDGRPLITI-----PSHTKQVL--AVTFSPDQGTIVSAGADNTVKLW 1297

QY 325 DIHTSGLLGEIHTGHSTIQYCDPSPQNHVALVSOYCVLWNTDSRSKVADCRGLSW 884
DB 1298 S--BNGTLTTLLEGNEAVQVIFSPDGRLIATASADKTIITLWSRDG-NILGTAFAGNHE 1354

QY 885 VHGVNFPDGSFSTSSDDQTIRLWETKVKCKNSAVMLKQEVVVFOENE---VMVLAVD 941
DB 1355 VNSLSFDPDGNILASGDDNTVRLWTVNRTLPKTFYGHKGSVSVYRFPNDGKKTSLSTD 1414

QY 942 HIRBLQINGRTGQIDYLTEAQVSCCLSPHLOVIAFGDENGAIIBELVNNR----IPQS 998
DB 1415 STMTKWSLDGKLLQTLSSPLPDVTSISFTPDNKIVALASPDHTIHLV---NRQGLLRS 1470

QY 999 RFQHKVTVHIOPTADEKTLISSDDAEIQW--NWQLDKCIFLRGHQETVKDFPLLNK 1056
DB 1471 LPGNHHTITSLSPFNKQILASGADKTIKLSVNGRLKT--LLGHNGWVTDIKFSADG 1528

QY 1057 R-LLSWSPDGTGVKVNIIITGNKEKDFVCHQCTVLSCDISHDATKFSSTSAKTAKIWSFD 1115
DB 1529 KNIYASADTKIWS-LDGRLIIRTLQHSASVSNVLSPPGOTLASTSQDETIKLWNL 1587

QY 1116 LLLPHELHRLHNGCVRCSAFSDVSTLLATGDNGEIRIWNYSNGELLHLCAPLSEGAAT 1175
DB 1588 GEL-IYTLRGHSDVVNLSFSPDCKTIASASDDGTIKLWVNPNGTLK-----TFQG--- 1638

QY 1176 HGGWVTDLCRSPDCKMLISAGY---IKWNV 1204
DB 1639 HRGGRSVSPDCKILAS-GGHTTVKVNWL 1669

RESULT 14
YL24 ANASP
ID YL24 ANASP STANDARD; PRG: 1683 AA.
AC Q8YV57;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical WD-repeat protein all2124.
GN OrderedLocusNames=all2124;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
NCBI_TaxID=103690;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yanada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
CC -1- SIMILARITY: Contains 13 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP003588; BAB73823.1; -.
DR PIR; AF2071; AF2071.
DR HSP; P16649; 1ERJ.
DR InterPro; IPR004106; Peptidase S9A.N.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR011680; WD40.
DR Pfam; PF00400; WD40; 14.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00320; WD40; 14.
DR PROSITE; PS00678; WD_REPEATS_1; 3.
DR PROSITE; PS50082; WD_REPEATS_2; 14.

DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Complete proteome; Hypothetical protein; Repeat; WD repeat.
FT REPEAT 1070 1109 WD 1.
FT REPEAT 1111 1150 WD 2.
FT REPEAT 1153 1192 WD 3.
FT REPEAT 1195 1236 WD 4.
FT REPEAT 1280 1318 WD 5.
FT REPEAT 1320 1359 WD 6.
FT REPEAT 1361 1400 WD 7.
FT REPEAT 1402 1441 WD 8.
FT REPEAT 1444 1483 WD 9.
FT REPEAT 1486 1525 WD 10.
FT REPEAT 1528 1567 WD 11.
FT REPEAT 1570 1609 WD 12.
FT REPEAT 1611 1650 WD 13.
SQ SEQUENCE 1683 AA; 185459 MW; F7CB361F547137 CRC64;
Query Match 10.2%; Score 672; DB 1; Length 1683;
Best Local Similarity 23.2%; Pred. No. 5.1e-30;
Matches 281; Conservative 196; Mismatches 490; Indels 244; Gaps 41;
QY 95 SSSGKDSVSGTSTVTVLCEGVQRPVVFTRKLVNAIQKLSKLGEPEGWTVIHG 154
DB 553 ASGSGKSS-----VVOAGLIPQ-----LRQGHIPNSEQWIKTIR--PQVNPLEA 596
QY 155 MACGKSVLAFAVRDHSLLPGCF-----PGVHWSVGKQKSGLLM-----KLQNLCTRL 206
DB 597 LAR-----KLGEWGTHLLIEGMLHOGVEGVVYLRSLPQGVTVLVIDQFEELFTLAPT 651
QY 207 DQDESQRLPNTIEAKRLRLILMKRKHPRSLILLDDVDSWVLKAFDSQCQILLTTRD 266
DB 652 DR-ELFLELLGAVQAGDFKLII-----TLRA----- 679
QY 267 KSVTDSVMGPKVYVSVSSLGKGLIEILSLFVNKKADLPQOASIIKECK--GSPLVV 324
DB 680 -----DFAIPCEVPAALQVAVSLV--PEKLSLDYRRVLHPAQVGLQVERA 728
QY 325 SLIGALLR-----DPNRWEYVLQLOKQKFKRKSSVD----- 360
DB 729 ELVEVLRLNQSVCGLP-LLEFVLEQLAQRTAGKLTLSYQEQLOGKIGALERSQGV 787
QY 361 YEALDEAMS-----TSVEMLEDIKD-----YVTDL-----SILQKQVVPKVL 401
DB 788 YESLPQLOECNAWIFSLTQLEGIEDFRRIYSDLIIVKYFAGLVQETLNLVNAKL 847
QY 402 ILWDMETEVEDILOEFVNKSLFLCDRNGKSPRYVLDQVDFLTERKNCSQLQDLHKII 461
DB 848 VVINLEAE-----IEAQGKSFSTPPNLSFTFVVEVAHEILIRHWSTL 891
QY 462 TQQRVHQSHTLSPPQ--EDCMVY-----NPLAYHMASAMHKELCALMFSLDWIKAK 513
DB 892 RNLWLENDRLRKQRQINHACQLWQSGQKQADFL---LQGARLAABEDIYIYTDDELGAD 948
QY 514 T-ELVGP--AHLIHFVEYRHIIDEKDCAV-----SENFOEFL 548
DB 949 VQEFITGACLAERKIQOLQAKNRKQARAVALSIVLGIASVSFGGLAYWQGREAQFRETA 1008
QY 549 SLN-----GHLLGRQFPFNIVQLGCEPETSVEVQAKLQAKQEVNDGMLVLEWINKNIT 604
DB 1009 ALNSSSSQANLLSHQOLAALI-----ASLKAQVQNVHVIAPNNKLAVTV 1053
QY 605 NLSRLV-----VRPHDAVYHACFSDGQRIASCADKTLQVFAEYKGLLEIKAK 656
DB 1054 TLQOALFEMQERNRLEHKGDVISISISRDGQTIASGSLDKTIKLMSRD--GRLFRTLNHG 1112
QY 657 EDEVLCCAFSTDTRFIATCSVDKVKIWNMTGELVHYVDHSHRQVNCCHFTNSHLLH 716
DB 1113 EDVYVSFSFGQTIASGSDKTIKLQTSQDGLTKITGHEQTVNNVVFSPDGN--L 1170
QY 717 ATGSSDCEKLKLDLWLNQKCRNTMFTGHTNSVNHCRSPDDKLLASCADGTLKLDWATSAN 776
DB 1171 ASASSDHSIKLNDTTSQGLLMTLTGHSAGVIIVRSPDQTIAGSBDKTVKLWHRQDK 1230

QY 777 ERKSINQKQFFLANLEPOEDMEVIVKCCWSADGARINVAANKKIFLFDIHTSGLIGEIH 836
DB 1231 LLKTLNGHQDWN-----SLSFSPDGKTLASASADKTIKLMRIADGKLVKTL 1277
QY 837 TGHSTIOYCDFPSQNHIAVVALSVQCVELMNTDSRSKVDRCRCHLSWVHGVWFSPDGSS 896
DB 1278 KGHNSDWDVNFSSDGGKATASARDNTIKLMNRHG--IELETFTHSGSGVAVVNFPLDSNI 1336
QY 897 FLTSSDDOTRLWE-----TKVKCKNSAVMLKQEVDFVQENEVWVLAVDHRLRLQLIN 950
DB 1337 TASASLNDTILMORPLISPLEVLAGNSGVYA-----VSPLHDSIIATAGADGNIQLWH 1391
QY 951 GRTGQIDYLTAEQVSCCLLS--PHLOYTAFDENGAIETILELVNRRIFQSRFOHKKTVWH 1008
DB 1392 SQDGLSKLTPGKAIYVIGFTFQDGLIASANADKTVKIWRVDRGKALKTLIGHDNEVNK 1451
QY 1009 TOFTADEKTLISSDDAEIOVNWQDKCIF---LRHQSTVVKDFRLKNSRL--SWSFD 1064
DB 1452 VNFSPDGKTLASARDNTVKLMNVSDGK--FKCTLKGTDEVFWVFSFGDKIIASASAD 1509
QY 1065 GTVKVMNIITGNKEKDFVCHQGTVLSCDISHDATKFSSTSDAKTIW--SPDLLPLHEL 1123
DB 1510 KTIRLMDSFGNLIKSLPAHNDLVYSVNFDPDGMASLASTGADTKVLRSHDGHIL--LHTF 1568
QY 1124 RHNGCVRCSAFVSVDSTLLATGDNDGGEIRIWNVNGELLHLCAPLSEGAATHGGWVTDL 1183
DB 1569 SGHNVVYSSFSFGDGRYIASASEDKTVKIWDG----HLLTTLPO-----HQAGVMSA 1619
QY 1184 CFSPDGKMLIS 1194
DB 1620 IFSPDGKTLIS 1630
RESULT 15
Q7ND05 PRELIMINARY; PRT; 1193 AA.
AC Q7ND05;
DT 01-VAR-2004 (Tremblrel. 26, Created)
DT 01-VAR-2004 (Tremblrel. 26, Last sequence update)
DT 01-VAR-2004 (Tremblrel. 26, Last annotation update)
DE WD-repeat protein.
GN OrderedocushNames=glr2821;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
CX NCBI_TaxID=33072;
RP [1]
SEQUENCE FROM N.A.
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003).
CC -!- SIMILARITY: Contains 14 WD repeats.
EMBL; AP006578; BAC90762.1; --
GO; GO:0042829; P:defense response to pathogen; IEA.
DR InterPro; IPR000767; Disease_resist.
DR Pfam; PF00400; WD40. 14.
DR PRINTS; PR00364; DISEASERSIST.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD00018; WD40. 13
DR PROSITE; PS00678; WD_REPEATS_1; 11.
DR PROSITE; PS00892; WD_REPEATS_2; 14.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Complete proteome; Repeat; WD-repeat;
SQ SEQUENCE 1193 AA; 130908 MW; 07AB8750D470E6C3 CRC64;
Query Match 10.1%; Score 666; DB 2; Length 1193;
Best Local Similarity 23.1%; Pred. No. 6.9e-30;



GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 18:25:54 ; Search time 176 Seconds
(without alignments)
2543.717 Million cell updates/sec

Title: US-10-646-396-2

Perfect score: 6619

Sequence: 1 MDKARNCLLQHREALEKDI.....FKTVTVVDNLGILYILQTL 1248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	6619	100.0	1248	4	AAy97636
2	6619	100.0	1248	5	ABG65661
3	6619	100.0	1248	6	ABP72163
4	6616	100.0	1248	4	AAy97646
5	6616	100.0	1248	4	AAy97639
6	6616	100.0	1248	4	AAy97643
7	6615	99.9	1248	4	AAy97644
8	6614	99.9	1248	4	AAy97645
9	6614	99.9	1248	4	AAy97641
10	6613	99.9	1248	4	AAy97647
11	6611	99.9	1248	4	AAy97642
12	6548.5	98.9	1237	4	AAy97638
13	6548.5	98.9	1237	7	ADP65308
14	6352.5	96.0	1205	2	AAW91072
15	6352.5	96.0	1205	4	AAy97637
16	6282	94.9	1194	2	AAW91071
17	6282	94.9	1194	3	AAy49432
18	6282	94.9	1194	4	AAy97649
19	6282	94.9	1194	5	ABG65662
20	6282	94.9	1194	6	ABB82741
21	6282	94.9	1194	7	ADP65303
22	6282	94.9	1194	8	ADG87139
23	6112	92.3	1149	4	AAy97640
24	6003.5	90.7	1199	8	ADM87180
25	5933.5	89.6	1526	8	ADM87179

26	5817	87.9	1238	5	ABG65660	Abg65660	Mouse apo
27	3205.5	48.4	769	4	ABG13410	Abg13410	Novel hum
28	2977	45.0	570	4	AAy97648	AAy97648	Apaf-1XL(
29	1589	24.0	308	5	ABJ04764	Abj04764	Apaf-1 pr
30	492	7.4	97	5	ABB81757	Abb81757	Tumour ne
31	476	7.2	93	5	ABJ04761	Abj04761	Apaf-1 pr
32	423	6.4	2629	2	AAW55885	Aaw55885	Rat telom
33	423	6.4	2629	7	ADD21418	Add21418	Rat TERT
34	423	6.4	2629	7	ADD21417	Add21417	Rat TERT
35	423	6.4	2629	7	ADD21424	Add21424	Mouse TER
36	423	6.4	2629	7	ADE63119	Ade63119	Rat Prote
37	413	6.2	2627	7	ADE63121	Ade63121	Human Pro
38	410.5	6.2	2627	2	AAW61347	Aaw61347	Human tel
39	410	6.2	2630	7	ADK40934	Adk40934	Novel hum
40	405	6.1	2625	2	AAW55887	Aaw55887	Human tel
41	399.5	6.0	2629	2	AAW61348	Aaw61348	Mouse tel
42	346	5.2	341	3	AAG38744	Aag38744	Arabidops
43	342.5	5.2	361	4	ABB68576	Abb68576	Drosophil
44	337.5	5.1	520	5	ABP73383	Abp73383	Candida a
45	337	5.1	797	5	ABP73371	Abp73371	Candida a

ALIGNMENTS

RESULT 1
AAy97636
ID AAy97636 standard; protein; 1248 AA.
XX
AC AAy97636;
XX
DT 20-APR-2001 (first entry)
XX
DE Apaf-1XL protein sequence.
XX
KW Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
KW hereditary disease; Apaf-1XL.
XX
OS Homo sapiens.
XX
PN WO200100827-A1.
XX
PD 04-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US018039.
XX
PR 30-JUN-1999; 99US-014718P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Nunez G, Hu Y;
XX
DR WPI; 2001-112454/12.
DR N-PSDB; AAA91115.
XX
PT Compositions for screening apoptosis pathway agonists and antagonists,
PT useful for the treatment and diagnosis of cancer and autoimmune diseases,
PT comprises new splice variants of wild-type Apaf-1 gene.
XX
PS Disclosure; Fig 25; 101pp; English.
XX
CC This sequence represents an Apaf-1 variant of the invention. The variants
CC of the invention are all splice variants of a wild-type Apaf-1 gene (a
CC mammalian homolog of CED-4 that participates in cytochrome c-dependent
CC activation of caspase-3). The DNA is useful in screening assays for
CC identifying apoptosis signalling pathway (antagonists, which are in turn
CC useful as potential therapeutic and diagnostic or prognostic tools for
CC diverse types of cancers, autoimmune diseases and hereditary diseases and
CC for screening compounds that modulates the interaction of Apaf-1 with
CC other members of the signalling pathway, i.e. their substrates or
CC ligands. The protein encoded by the Apaf-1 variant genes are useful in a
CC cell-free assay system and the antibody generated to the translation

CC tumour formation. This is the amino acid sequence of an Apaf-1 protein
 CC DNA encoding which used in the creation of antisense oligonucleotides for
 CC modulating the activity and expression of Apaf-1
 XX
 SQ Sequence 1248 AA;

Query Match	100.0%;	Score 6619;	DB 5;	Length 1248;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 1248;	Conservative 0;			

Qy	1	MDAKARNCLLQHREALEKDIKTSYIMDMHISDGLFTISEEEKVNRNEPTQOQRAAMLKMI	60
Db	1	MDAKARNCLLQHREALEKDIKTSYIMDMHISDGLFTISEEEKVNRNEPTQOQRAAMLKMI	60
Qy	61	LKKNDSTVSYFNALLHGGYKDLAALLHDGIPVSSSSGKDSVSGITSYRTVLCGGVP	120
Db	61	LKKNDSTVSYFNALLHGGYKDLAALLHDGIPVSSSSGKDSVSGITSYRTVLCGGVP	120
Qy	121	QRPVVVTRKKLVNAIQKLSKLGEPGWVTHGAGCGKSVLAABAARDHSLLEGCPG	180
Db	121	QRPVVVTRKKLVNAIQKLSKLGEPGWVTHGAGCGKSVLAABAARDHSLLEGCPG	180
Qy	181	GVHVSVGKQKSGLLMKLQNLCTRLDQESFQSLPLNIEEAKDRLRLMLRKHPRSL	240
Db	181	GVHVSVGKQKSGLLMKLQNLCTRLDQESFQSLPLNIEEAKDRLRLMLRKHPRSL	240
Qy	241	ILDDVMDSWLKAQFSDQCOILLTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN	300
Db	241	ILDDVMDSWLKAQFSDQCOILLTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN	300
Qy	301	MKKADLPEQAHSIKECKGSLVSLIGALLRDPNRPWEYVLKQONKQFRIKSSSYD	360
Db	301	MKKADLPEQAHSIKECKGSLVSLIGALLRDPNRPWEYVLKQONKQFRIKSSSYD	360
Qy	361	YEALDEAMSISVEMLRDIDKYITDLSLQKDVKPTKVLCLMDMETEEVEDILOQFVN	420
Db	361	YEALDEAMSISVEMLRDIDKYITDLSLQKDVKPTKVLCLMDMETEEVEDILOQFVN	420
Qy	421	KSLLFCDRNGKSPRYLHDLQVDFLTKNCSQLODLHKKIIITQFQYHQPHLSPDQEDC	480
Db	421	KSLLFCDRNGKSPRYLHDLQVDFLTKNCSQLODLHKKIIITQFQYHQPHLSPDQEDC	480
Qy	481	MYWTFNFLAYHMASAKMHELCALMPSLDWIKAETELVGPALHIEFVEYRHLDEKCAV	540
Db	481	MYWTFNFLAYHMASAKMHELCALMPSLDWIKAETELVGPALHIEFVEYRHLDEKCAV	540
Qy	541	SENFQEFSLNGLHLLGRQPFNIVOLGLCEPETSEVYQOAKLOAKQEVNDGMLYLEWINK	600
Db	541	SENFQEFSLNGLHLLGRQPFNIVOLGLCEPETSEVYQOAKLOAKQEVNDGMLYLEWINK	600
Qy	601	KNITNLSRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFAETGEKLLLEIKAHEDV	660
Db	601	KNITNLSRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFAETGEKLLLEIKAHEDV	660
Qy	661	LCCAFSTDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS	720
Db	661	LCCAFSTDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS	720
Qy	721	SDCFKLKLDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS	780
Db	721	SDCFKLKLDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS	780
Qy	781	INVQKFFLNLPDQEDMEVIVKCCSWSDAGGARIMVAANKIFLFDIHTSGLLGEIHTGHH	840
Db	781	INVQKFFLNLPDQEDMEVIVKCCSWSDAGGARIMVAANKIFLFDIHTSGLLGEIHTGHH	840
Qy	841	STIQYCDSPQNHLLAVALSQYVELMNTDTSRKSVDRCRHLWSVHGVMPDSSSFLTS	900
Db	841	STIQYCDSPQNHLLAVALSQYVELMNTDTSRKSVDRCRHLWSVHGVMPDSSSFLTS	900
Qy	901	SDDOTIRLWETKVKCKSAVMLEKQEVVDVFOENEMVLAVDHIRLQINGRTGQIDYLT	960
Db	901	SDDOTIRLWETKVKCKSAVMLEKQEVVDVFOENEMVLAVDHIRLQINGRTGQIDYLT	960

Qy	961	EAQVSCCCLSPHLQVIAFGDENGAIIELELVNRRIFQSRPQHKKTVMHIQFTADEKTLIS	1020
Db	961	EAQVSCCCLSPHLQVIAFGDENGAIIELELVNRRIFQSRPQHKKTVMHIQFTADEKTLIS	1020
Qy	1021	SSDDAEIOVWNWQDKCIFLRGHQETVKDFRLLKNSRLLSWSFDGTVKVMNIITGNKEK	1080
Db	1021	SSDDAEIOVWNWQDKCIFLRGHQETVKDFRLLKNSRLLSWSFDGTVKVMNIITGNKEK	1080
Qy	1081	FVCHQGTVLSCDISHDATKFSSTADTKIWSFDLLPLHLRHGNGCVCSAFSVDST	1140
Db	1081	FVCHQGTVLSCDISHDATKFSSTADTKIWSFDLLPLHLRHGNGCVCSAFSVDST	1140
Qy	1141	LLATGDDNGEIRIWNVNGELHLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGYIK	1200
Db	1141	LLATGDDNGEIRIWNVNGELHLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGYIK	1200
Qy	1201	WNVVTGESSOTFTYNTGNTLAKIHSVDPFKTVTVTDNLGILYILOTLE	1248
Db	1201	WNVVTGESSOTFTYNTGNTLAKIHSVDPFKTVTVTDNLGILYILOTLE	1248

RESULT 3

ID	ABP72163	standard; protein; 1248 AA.
XX	ABP72163;	
AC	AC	
XX	XX	
DT	22-APR-2003	(first entry)
XX	XX	
DE	Human APAF1, inducer of cell death.	
XX	XX	
KW	Human; APAF1; cell death; apoptosis; neurodegenerative disease;	
KW	heart disease; cardiomyopathy; cardiant; neuroprotective; gene therapy;	
KW	gene; 'ss.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
PN	WO2003004606-A2.	
XX	XX	
PD	16-JAN-2003.	
XX	XX	
PF	03-JUL-2002; 2002WO-US021002.	
XX	XX	
PR	03-JUL-2001; 2001US-00898158.	
XX	XX	
PA	(UYCO) UNIV COLUMBIA NEW YORK.	
XX	XX	
PI	Troy CW, Shelanski ML;	
XX	XX	
DR	WPI; 2003-210351/20.	
DR	N-PSDB; ABZ58107.	
XX	XX	
PT	New nucleic acid encoding an inhibitor-of-apoptosis protein, useful for	
PT	treating cancer, neurodegenerative disorder or cardiomyopathy.	
PS	Disclosure; Fig 21A; 124pp; English.	
XX	XX	
CC	The present sequence is the protein sequence for human APAF1, a protein	
CC	that induces cell death. The invention provides a nucleic acid, such as	
CC	an antisense oligonucleotide, which specifically hybridises to a nucleic	
CC	acid encoding a protein that induces cell death, especially APAF1, RAIDD	
CC	or Diablo/SMAC. A claimed method for inhibiting a cell's death	
CC	(especially a neuronal cell's death) comprises contacting the cell with	
CC	the nucleic acid under conditions permitting the nucleic acid to enter	
CC	the cell, especially the use of a vector, liposome, or a mechanical or	
CC	electrical means. The method is used to treat a neurodegenerative	
CC	disorder, especially a brain disorder or central nervous system disorder,	
CC	or a heart disorder, especially cardiomyopathy, in a human (all claimed)	
XX	XX	
SQ	Sequence 1248 AA;	

Query Match 100.0%; Score 6619; DB 6; Length 1248;

		Best Local Similarity	100.0%;	Pred. No. 0;	Matches	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MDAKARNCLLOHREALEKDIKTSYIMDHMSIDGFLTISEEEKVNRNEPTQOQRAAMLKMI	60									
Db	1	MDAKARNCLLOHREALEKDIKTSYIMDHMSIDGFLTISEEEKVNRNEPTQOQRAAMLKMI	60									
Qy	61	LKXNDSDYVSFYNNALLHGGYKDLAALLHDGI PVVSSSSGKDSVSGITSYVTVLCEGVP	120									
Db	61	LKXNDSDYVSFYNNALLHGGYKDLAALLHDGI PVVSSSSGKDSVSGITSYVTVLCEGVP	120									
Qy	121	QRPVVFTRKLVNAIQKLSKLBEPGWYTHGMAGCKSVLAARAVRSHSLEGCFCG	180									
Db	121	QRPVVFTRKLVNAIQKLSKLBEPGWYTHGMAGCKSVLAARAVRSHSLEGCFCG	180									
Qy	181	GVHWSVGKODKGLMLKQNLCTRLDQESFORLPINIEAKDRILMLRKHRSLL	240									
Db	181	GVHWSVGKODKGLMLKQNLCTRLDQESFORLPINIEAKDRILMLRKHRSLL	240									
Qy	241	ILDVDWDSVILKAFOSQOQIILLTRDKSVTDSVMGPKYVVPVSSSLGKGLIILSLFVN	300									
Db	241	ILDVDWDSVILKAFOSQOQIILLTRDKSVTDSVMGPKYVVPVSSSLGKGLIILSLFVN	300									
Qy	301	MKXADLPQOAHSIIEKCKGSLVLSLIGALLRFPNMEYILKQIQKQFKRIRKSSVD	360									
Db	301	MKXADLPQOAHSIIEKCKGSLVLSLIGALLRFPNMEYILKQIQKQFKRIRKSSVD	360									
Qy	361	YEALDEAMSISVEMLRDIDKYITDLSILOKQVKVPTKVLICILMDMETEEVEDILOEFVN	420									
Db	361	YEALDEAMSISVEMLRDIDKYITDLSILOKQVKVPTKVLICILMDMETEEVEDILOEFVN	420									
Qy	421	KSLFLCDNRNGKFRYLLHDLQVDFLTKNCSQDLHKKIITQFORVHQPHLSPQEDC	480									
Db	421	KSLFLCDNRNGKFRYLLHDLQVDFLTKNCSQDLHKKIITQFORVHQPHLSPQEDC	480									
Qy	481	MYWYNFLAYHMASAKWHELCALMFSLDWIKATELVGPAHLIHEFVEYRHLDEKCAV	540									
Db	481	MYWYNFLAYHMASAKWHELCALMFSLDWIKATELVGPAHLIHEFVEYRHLDEKCAV	540									
Qy	541	SENFOEFLSLNGLLGRPPENIVQLGLCEPTESEVYQAKIQAKQEVNDGMLYLEWINK	600									
Db	541	SENFOEFLSLNGLLGRPPENIVQLGLCEPTESEVYQAKIQAKQEVNDGMLYLEWINK	600									
Qy	601	KNITNLSRLVVRPHTDAVYHACSEDOGRIASCGADKTLQVFKAEKGKLEIKAHEDV	660									
Db	601	KNITNLSRLVVRPHTDAVYHACSEDOGRIASCGADKTLQVFKAEKGKLEIKAHEDV	660									
Qy	661	LCCAFSTDRFATCSVDKKIWNSTGELVHTYDEHSEQVNCCHFTNSHLLATGS	720									
Db	661	LCCAFSTDRFATCSVDKKIWNSTGELVHTYDEHSEQVNCCHFTNSHLLATGS	720									
Qy	721	SDCFLKMLDQKCRNMTFHTNSVNHCRFPDDKLLASCADGTLKIMDATSANERKS	780									
Db	721	SDCFLKMLDQKCRNMTFHTNSVNHCRFPDDKLLASCADGTLKIMDATSANERKS	780									
Qy	781	INVKQFFNLNEDPQEDMEVIVKCCSADGARIMVAANKIFLFDIHTSGLLGEIHTGHH	840									
Db	781	INVKQFFNLNEDPQEDMEVIVKCCSADGARIMVAANKIFLFDIHTSGLLGEIHTGHH	840									
Qy	841	STIQYCDSPQHLAVVALSOYCVELWNTDSRSKVADCRGHLVHVGWVMSFGSSFLTS	900									
Db	841	STIQYCDSPQHLAVVALSOYCVELWNTDSRSKVADCRGHLVHVGWVMSFGSSFLTS	900									
Qy	901	SDQITRLWETKVKCKSAVMKQEVVDVVFQENWVLAVDHRRLOLINGRTGQIDYLT	960									
Db	901	SDQITRLWETKVKCKSAVMKQEVVDVVFQENWVLAVDHRRLOLINGRTGQIDYLT	960									
Qy	961	EAQVSCCLSLPHLOVIAFGDENGAIETLELVNRRIFQSRFOHKKTVWHIOTFADEKTLIS	1020									
Db	961	EAQVSCCLSLPHLOVIAFGDENGAIETLELVNRRIFQSRFOHKKTVWHIOTFADEKTLIS	1020									
Qy	1021	SSDDAEIQVNNWQDKCIFLRGHQETVKDFRLLKNSRLLSWSPDGTVKVNNIITGNKED	1080									

Db	1021	SSDDAEIQVNNWQDKCIFLRGHQETVKDFRLLKNSRLLSWSPDGTVKVNNIITGNKEK	1080
Qy	1081	FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLPLHLRGNHGCVRCSAFSDST	1140
Db	1081	FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLPLHLRGNHGCVRCSAFSDST	1140
Qy	1141	LLATGDDNGEIRIWNVNSGELLHLCAPLSEGAATHGCGWTDLCFSPDGKMLISAGYIK	1200
Db	1141	LLATGDDNGEIRIWNVNSGELLHLCAPLSEGAATHGCGWTDLCFSPDGKMLISAGYIK	1200
Qy	1201	WNVVTVGESSQTFYNGTNLKKIHVSPDKFTVTVVDNLGILYILQITLE	1248
Db	1201	WNVVTVGESSQTFYNGTNLKKIHVSPDKFTVTVVDNLGILYILQITLE	1248
RESULT 4			
ID	AA97646	standard; protein; 1248 AA.	
XX	AA97646;		
DT	20-APR-2001	(first entry)	
DE	Apaf-1XL-Kl60R	protein sequence.	
XX	Apaf1; CED-4	homologue; cytochrome c-dependent activation; caspase-3;	
XX	apoptosis signalling pathway; cancer; autoimmune disease; variant;		
XX	hereditary disease; Apaf-1XL-Kl60R.		
OS	Homo sapiens.		
XX	WO200100827-A1.		
XX	04-JAN-2001.		
XX	30-JUN-2000;	2000MO-US018039.	
XX	30-JUN-1999;	99US-0141718P.	
XX	(UNMI)	UNIV MICHIGAN.	
PI	Nunez G, Hu Y;		
DR	WPI; 2001-112454/12.		
DR	N-PSDB; AAA91125.		
PT	Compositions for screening apoptosis pathway agonists and antagonists, useful for the treatment and diagnosis of cancer and autoimmune diseases, comprises new splice variants of wild-type Apaf-1 gene.		
PS	Disclosure; Fig 35; 101pp; English.		
XX	This sequence represents an Apaf-1 variant of the invention. The variants of the invention are all splice variants of a wild-type Apaf-1 gene (a mammalian homolog of CED-4 that participates in cytochrome c-dependent activation of caspase-3). The DNA is useful in screening assays for identifying apoptosis signalling pathway (ant)agonists, which are in turn useful as potential therapeutics and diagnostic or prognostic tools for diverse types of cancers, autoimmune diseases and hereditary diseases and for screening compounds that modulates the interaction of Apaf-1 with other members of the signalling pathway, i.e. their substrates or ligands. The protein encoded by the Apaf-1 variant genes are useful in a cell-free assay system and the antibody generated to the translation product are used in immunoprecipitation assays to isolate new Apaf-1 pathway constituents or their natural mutants		
SQ	Sequence 1248 AA;		
Query Match 100.0%; Score 6616; DB 4; Length 1248;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1247; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Ov	1	MDAKARNCLLOHREALEKDIKTSYIMDHMSIDGFLTISEEEKVNRNEPTQOQRAAMLKMI	60

Db 1 MDKARNCLLQHREALEKDIKTSYIMDMHISDGLFTTISEEEKVRNEPTQOORAAALIKMI 60
Qy 61 LKNDSDSVSYFNALLHGGYKDLAALLHGGIPVWSSSGKDSVSGITSYVTVLCEGVP 120
Db 61 LKNDSDSVSYFNALLHGGYKDLAALLHGGIPVWSSSGKDSVSGITSYVTVLCEGVP 120
Qy 121 QRPVVVTRKLVNAIQKLSKLGEPGWVTHGMAGCGKSVLAABAVRDSHSLLEGCPG 180
Db 121 QRPVVVTRKLVNAIQKLSKLGEPGWVTHGMAGCGKSVLAABAVRDSHSLLEGCPG 180
Qy 181 GVHVSVCKQKSLGLMLKQNLCTELDODESFSQRLPLNIEAKDRILMLRKHPRSL 240
Db 181 GVHVSVCKQKSLGLMLKQNLCTELDODESFSQRLPLNIEAKDRILMLRKHPRSL 240
Qy 241 ILDDVWDSWLKAFDSQCOILLTRDKSVTSVMGPKYVVPVSSSLGKEGLEILSLFN 300
Db 241 ILDDVWDSWLKAFDSQCOILLTRDKSVTSVMGPKYVVPVSSSLGKEGLEILSLFN 300
Qy 301 MKKADLPQAHSIIEKCGSLPWSLIGALLRDPFNRWEYVYLKQNLQKQFRIKSSSYD 360
Db 301 MKKADLPQAHSIIEKCGSLPWSLIGALLRDPFNRWEYVYLKQNLQKQFRIKSSSYD 360
Qy 361 YEALDEAMSISVEMLRDIDKYYTDLSTLOKDVKPTKVLCLMDMETEEVEDILOEFVN 420
Db 361 YEALDEAMSISVEMLRDIDKYYTDLSTLOKDVKPTKVLCLMDMETEEVEDILOEFVN 420
Qy 421 KSLFLCDNRNGKSFYRLHDLQVDFTEKNCQLODLHKKIITQORYHOPHTLSPDQDC 480
Db 421 KSLFLCDNRNGKSFYRLHDLQVDFTEKNCQLODLHKKIITQORYHOPHTLSPDQDC 480
Qy 481 MYWNFLAYHMASAKMKELCALMFLSDWIKAKTELVGPAHLIHEFVEYRHLDEKCAV 540
Db 481 MYWNFLAYHMASAKMKELCALMFLSDWIKAKTELVGPAHLIHEFVEYRHLDEKCAV 540
Qy 541 SENQEFSLNGHLIGROPFPNIIVOLGLCEPETSEVYQOAKLOAQKQEVNDGMLYLEWINK 600
Db 541 SENQEFSLNGHLIGROPFPNIIVOLGLCEPETSEVYQOAKLOAQKQEVNDGMLYLEWINK 600
Qy 601 KNIITNLSLVVRPHTDAYVHACFSDGQRIASCADKTLQVFKAGTGEKLEIKAHDEV 660
Db 601 KNIITNLSLVVRPHTDAYVHACFSDGQRIASCADKTLQVFKAGTGEKLEIKAHDEV 660
Qy 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
Db 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
Qy 721 SDCFLKLDLWLNQKCRNTMFGHTSVNHCPSDDKLLASCADGTLKLDATSANERKS 780
Db 721 SDCFLKLDLWLNQKCRNTMFGHTSVNHCPSDDKLLASCADGTLKLDATSANERKS 780
Qy 781 INVQOFFLNLSDPOEDMEVIVKCCSWASDAGRIWVAANKIFLFDIHTSGLLGEIHTGHH 840
Db 781 INVQOFFLNLSDPOEDMEVIVKCCSWASDAGRIWVAANKIFLFDIHTSGLLGEIHTGHH 840
Qy 841 STIQYCDSPQNLAVVALSQCVELMNTDSRSKVADCRGHLVHGVWVSPDSSFLTS 900
Db 841 STIQYCDSPQNLAVVALSQCVELMNTDSRSKVADCRGHLVHGVWVSPDSSFLTS 900
Qy 901 SDDQIRLWETKVKCNKSNVLMKEVDVVFQENVMVLAVDHIRRLQINRGTOIDYLT 960
Db 901 SDDQIRLWETKVKCNKSNVLMKEVDVVFQENVMVLAVDHIRRLQINRGTOIDYLT 960
Qy 961 EAQVSCCLSPHLOVIAFGDNGAILEILEVNNRIPOSFQHKTVWHIOFTADEKTLIS 1020
Db 961 EAQVSCCLSPHLOVIAFGDNGAILEILEVNNRIPOSFQHKTVWHIOFTADEKTLIS 1020
Qy 1021 SSDDAEIOVNWQDLKCIIFLAGHQBTVKDFRLLKNSRLLSFDCGTVKVWNIITGNKEK 1080
Db 1021 SSDDAEIOVNWQDLKCIIFLAGHQBTVKDFRLLKNSRLLSFDCGTVKVWNIITGNKEK 1080
Qy 1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLLPLHLRHGNCVRCASFVDS 1140

Db 1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLLLPLHLRHGNCVRCASFVDS 1140
Qy 1141 LLATGDDNGEIRIWNVNGELLHLCAPLSEGAATHGCGWTDLCFSPDGKMLISAGGYIK 1200
Db 1141 LLATGDDNGEIRIWNVNGELLHLCAPLSEGAATHGCGWTDLCFSPDGKMLISAGGYIK 1200
Qy 1201 WNNVVTGESSQTFYTNCTNGLKIHVSPOPKTYVTVVDNLGILYILQTL 1248
Db 1201 WNNVVTGESSQTFYTNCTNGLKIHVSPOPKTYVTVVDNLGILYILQTL 1248

RESULT 5
AA97639
ID AA97639 standard; protein; 1248 AA.
XX AA97639;
XX 20-APR-2001 (first entry)
XX Apaf-1XL-M368L protein sequence.
DE Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
KW hereditary disease; Apaf-1XL-M368L.
XX Homo sapiens.
XX WO200100827-A1.
XX 04-JAN-2001.
XX 30-JUN-2000; 2000WO-US018039.
XX 30-JUN-1999; 99US-0141718P.
XX (UNMI) UNIV MICHIGAN.
XX Nunez G, Hu Y;
XX WPI; 2001-112454/12.
XX N-PSDB; AAA91118.
XX Compositions for screening apoptosis pathway agonists and antagonists,
XX useful for the treatment and diagnosis of cancer and autoimmune diseases,
XX comprises new splice variants of wild-type Apaf-1 gene.
XX Disclosure; Fig 28; 101pp; English.
XX This sequence represents an Apaf-1 variant of the invention. The variants
XX of the invention are all splice variants of a wild-type Apaf-1 gene (a
XX mammalian homolog of CED-4 that participates in cytochrome c-dependent
XX activation of caspase-3). The DNA is useful in screening assays for
XX identifying apoptosis signalling pathway (antagonists, which are in turn
XX useful as potential therapeutics and diagnostic or prognostic tools for
XX diverse types of cancers, autoimmune diseases and hereditary diseases and
XX for screening compounds that modulates the interaction of Apaf-1 with
XX other members of the signalling pathway, i.e. their substrates or
XX ligands. The protein encoded by the Apaf-1 variant genes are useful in a
XX cell-free assay system and the antibody generated to the translation
XX product are used in immunoprecipitation assays to isolate new Apaf-1
XX pathway constituents or their natural mutants
XX Sequence 1248 AA;
XX

Query Match 100.0%; Score 6616; DB 4; Length 1248;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1247; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDKARNCLLQHREALEKDIKTSYIMDMHISDGLFTTISEEEKVRNEPTQOORAAALIKMI 60
Db 1 MDKARNCLLQHREALEKDIKTSYIMDMHISDGLFTTISEEEKVRNEPTQOORAAALIKMI 60
Qy 61 LKNDSDSVSYFNALLHGGYKDLAALLHGGIPVWSSSGKDSVSGITSYVTVLCEGVP 120

Db	61	LKONDSVSVFYALHHEGYKDLAALLHDLGIPVVSSSSKQSVSGITSYVTVWVLCSEGVV	120
Qy	121	QRPVFVTRKLVNA1QOKLSKLKGBEGWVTIHGMAGCGKSVLAABAVRHSLLLEGCPFG	180
Db	121	QRPVFVTRKLVNA1QOKLSKLKGBEGWVTIHGMAGCGKSVLAABAVRHSLLLEGCPFG	180
Qy	181	GVHWSVVGKDSGLLMKLQNLCTRLDQDSFSQRLPLNIEAKDRILRLMLRKHPSLL	240
Db	181	GVHWSVVGKDSGLLMKLQNLCTRLDQDSFSQRLPLNIEAKDRILRLMLRKHPSLL	240
Qy	241	ILDDVDSWVLKAFDSQCOILLTRDKSVSDTSVMGPKYVVPVSSSLGKXGLBILSLFVN	300
Db	241	ILDDVDSWVLKAFDSQCOILLTRDKSVSDTSVMGPKYVVPVSSSLGKXGLBILSLFVN	300
Qy	301	MKKADLPQOAHSI1KECKGSPLVWSLIGALLRDPFNKNWEYLLKOLQNKQFKRKSSSD	360
Db	301	MKKADLPQOAHSI1KECKGSPLVWSLIGALLRDPFNKNWEYLLKOLQNKQFKRKSSSD	360
Qy	361	YEALDZAMSASVEMLRREDIKDYVTDLS1LOKQVKVPKVLCLMDMTEERVEID1LOEPVN	420
Db	361	YEALDZAMSASVEMLRREDIKDYVTDLS1LOKQVKVPKVLCLMDMTEERVEID1LOEPVN	420
Qy	421	KSLFLCDNRNGKSPRYVLDLQVDFLTKNCSQLODLHKKITOPORVHOPHTLSPOEDC	480
Db	421	KSLFLCDNRNGKSPRYVLDLQVDFLTKNCSQLODLHKKITOPORVHOPHTLSPOEDC	480
Qy	481	MYWTFLAYHMASAKMHELCAIMFSLDWTKATELVGPAHLIHFVEYRHILDEKCAV	540
Db	481	MYWTFLAYHMASAKMHELCAIMFSLDWTKATELVGPAHLIHFVEYRHILDEKCAV	540
Qy	541	SENQPELSLNGHLGRQPPNI1VOLGLCBPETSSEVYQOAKLQAKQBVNDGMYLEWINK	600
Db	541	SENQPELSLNGHLGRQPPNI1VOLGLCBPETSSEVYQOAKLQAKQBVNDGMYLEWINK	600
Qy	601	KNITNLSRLVVRPHTDAVYHACFSEDQORIASCGADKTQVFKAETCEKLETKAHEDEV	660
Db	601	KNITNLSRLVVRPHTDAVYHACFSEDQORIASCGADKTQVFKAETCEKLETKAHEDEV	660
Qy	661	LCCAFSTDRTFATICSVDKKVKI1WNSMTGBLHVTYDEHSEQVNCCHPTNSHHLLIATGS	720
Db	661	LCCAFSTDRTFATICSVDKKVKI1WNSMTGBLHVTYDEHSEQVNCCHPTNSHHLLIATGS	720
Qy	721	SDCFLKLWD1NOKECRNTMFGHTNSVNHCRFSPDDKLLASCSADGTLKWDATASANERKS	780
Db	721	SDCFLKLWD1NOKECRNTMFGHTNSVNHCRFSPDDKLLASCSADGTLKWDATASANERKS	780
Qy	781	INVQFFLNLEDOEDMEVI1VKCCSWSADGAR1WVAANKI1FLFD1HTSGLLGEIHTGHH	840
Db	781	INVQFFLNLEDOEDMEVI1VKCCSWSADGAR1WVAANKI1FLFD1HTSGLLGEIHTGHH	840
Qy	841	STIQYCDSPQNHILAVALSQYCVELWN7DTSRKVADCRGHLSSVHGWFSPDGSFPTS	900
Db	841	STIQYCDSPQNHILAVALSQYCVELWN7DTSRKVADCRGHLSSVHGWFSPDGSFPTS	900
Qy	901	SDOQTELMETKKVCNSAVMLKQEDVVYFOENEVWLVADHTRBLQINGTGQIDVLT	960
Db	901	SDOQTELMETKKVCNSAVMLKQEDVVYFOENEVWLVADHTRBLQINGTGQIDVLT	960
Qy	961	EAQVSCCCLSPHLOVIATFDENGAEI1ELELVNRR1FQSRFOHKTTVWH1QFTADSEKTLIS	1020
Db	961	EAQVSCCCLSPHLOVIATFDENGAEI1ELELVNRR1FQSRFOHKTTVWH1QFTADSEKTLIS	1020
Qy	1021	SSDDAE1QVWNWOLDKCI1FLRGHQTETVKDFRLLKNSRLLSSWSPDGTWKVNI1ITGNKEK	1080
Db	1021	SSDDAE1QVWNWOLDKCI1FLRGHQTETVKDFRLLKNSRLLSSWSPDGTWKVNI1ITGNKEK	1080
Qy	1081	FVCHQGTVLSCDISHDATKFSSTSDAKTAKIWSFOLL1PLHELHGNGCVRCSAFSDVST	1140
Db	1081	FVCHQGTVLSCDISHDATKFSSTSDAKTAKIWSFOLL1PLHELHGNGCVRCSAFSDVST	1140
Qy	1141	LLATGDGNGEIR1WNVWSNGELLH1CAPLSEGAATHGQV7DLCFSPDGKMLISAGGYIK	1200

Db	1141	LLATGDDNGEIRINWVNSGELLHLCAPISEGAATHGGWTDLCFSPDGKMLISAGGYIK	1248
QY	1201	WNWVTGESSQFTYTGNTNLKIHVSFDFKTYVTVVDNLGILYILOTL	1248
Db	1201	WNWVTGESSQFTYTGNTNLKIHVSFDFKTYVTVVDNLGILYILOTL	1248
RESULT 6			
AA	AY97643		
ID	AA	AY97643 standard; protein; 1248 AA.	
XX	AA	AY97643;	
DT	20-APR-2001	(first entry)	
DE	Apaf-1XL-E39Q	protein sequence.	
KW	Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;		
KW	apoptosis signalling pathway; cancer; autoimmune disease; variant;		
KW	hereditary disease; Apaf-1XL-E39Q.		
OS	Homo sapiens.		
PN	WO200100827-A1.		
XX	04-JAN-2001.		
PF	30-JUN-2000; 2000WO-US018039.		
PR	30-JUN-1999; 99US-0141718P.		
PA	(UNMI) UNIV MICHIGAN.		
PI	Nunez G, Hu Y;		
DR	WPI; 2001-112454/12.		
XX	N-PSDB; AAA91122.		
PT	Compositions for screening apoptosis pathway agonists and antagonists,		
PT	useful for the treatment and diagnosis of cancer and autoimmune diseases,		
PT	comprises new splice variants of wild-type Apaf-1 gene.		
PS	Disclosure; Fig 32; 101pp; English.		
CC	This sequence represents an Apaf-1 variant of the invention. The variants		
CC	of the invention are all splice variants of a wild-type Apaf-1 gene (a		
CC	mammalian homolog of CED-4 that participates in cytochrome c-dependent		
CC	activation of caspase-3). The DNA is useful in screening assays for		
CC	identifying apoptosis signalling pathway (antagonists, which are in turn		
CC	useful as potential therapeutic and diagnostic or prognostic tools for		
CC	diverse types of cancers autoimmune diseases and hereditary diseases and		
CC	for screening compounds that modulates the interaction of Apaf-1 with		
CC	other members of the signalling pathway, i.e. their substrates or		
CC	ligands. The protein encoded by the Apaf-1 variant genes are useful in a		
CC	cell-free assay system and the antibody generated to the translation		
CC	product are used in immunoprecipitation assays to isolate new Apaf-1		
CC	pathway constituents or their natural mutants		
XX	Sequence 1248 AA;		
Query Match	100.0%; Score 6616; DB 4; Length 1248;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 1247; Conservative	1; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MDAKARNCLLQHREALEKIKTSYIMDMHISDGFYISEEKVRNEPTQQRAMLIKMI	60
Db	1	MDAKARNCLLQHREALEKIKTSYIMDMHISDGFYISEEKVRNEPTQQRAMLIKMI	60
QY	61	LKKDNDYSVSPYNALLHEGYKDLAALLHDGIPVYSSSGKDSVSGITSYVRTVLCGGVP	120
Db	61	LKKDNDYSVSPYNALLHEGYKDLAALLHDGIPVYSSSGKDSVSGITSYVRTVLCGGVP	120
QY	121	QRPVVFVTRKKLVNAIQOKLSKLKGEQWVTIHGMAGCGKSVLAAEAVRDHSLLLEGCFPG	180

Db 121 QRPVVVTRKKLVNAIQKLSKLGEPGWVTHGAGCGKSVLAEEAVRDSHSLLEGCPG 180
 Qy 181 GVHWSVGKQKSGLLMKLQNLCTRLDQDSEFSQRLPLNIEBEAKDRILMLRKHPSLL 240
 Db 181 GVHWSVGKQKSGLLMKLQNLCTRLDQDSEFSQRLPLNIEBEAKDRILMLRKHPSLL 240
 Qy 241 ILDDVWDSWVLFKAPDSQCIILTRDKSVTSVMGPKYVVPVSSSLGKEGLEILSLFVN 300
 Db 241 ILDDVWDSWVLFKAPDSQCIILTRDKSVTSVMGPKYVVPVSSSLGKEGLEILSLFVN 300
 Qy 301 MKKADLPQANSIIKECKGSLVLSLIGALLRDPENRWEYVYLKOLONKQKRIKSSSYD 360
 Db 301 MKKADLPQANSIIKECKGSLVLSLIGALLRDPENRWEYVYLKOLONKQKRIKSSSYD 360
 Qy 361 YEALDEAMSISVEMLREDIKDYTTDLISLOKDVKVPKVLCLMDMETEEVEDILQEFVN 420
 Db 361 YEALDEAMSISVEMLREDIKDYTTDLISLOKDVKVPKVLCLMDMETEEVEDILQEFVN 420
 Qy 421 KSLFLCDRNGKSFYYLHDLQVDFTEKNSQLOLHKKIITQORYHQPHTLSPDQEDC 480
 Db 421 KSLFLCDRNGKSFYYLHDLQVDFTEKNSQLOLHKKIITQORYHQPHTLSPDQEDC 480
 Qy 481 MYWYNFLAYHMASAKWHELCALMPSLOWIKAKTELVGPAHLIHEFVEYRHLDEKCAV 540
 Db 481 MYWYNFLAYHMASAKWHELCALMPSLOWIKAKTELVGPAHLIHEFVEYRHLDEKCAV 540
 Qy 541 SENFOEFLSLNGHLIGROFFNIVOLGICEPETSEVYQOAKLOAKQEVNDGMLYLEWINK 600
 Db 541 SENFOEFLSLNGHLIGROFFNIVOLGICEPETSEVYQOAKLOAKQEVNDGMLYLEWINK 600
 Qy 601 KNIITNLSRLVVRPHPTDAVYHAFCSBDGQRIASCGADKTLQVFKABTEGKLEIKAHEDV 660
 Db 601 KNIITNLSRLVVRPHPTDAVYHAFCSBDGQRIASCGADKTLQVFKABTEGKLEIKAHEDV 660
 Qy 661 LCCAFSTDRTATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGS 720
 Db 661 LCCAFSTDRTATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGS 720
 Qy 721 SDCFLKLDLWLNQECRNTMFGHTSNVHCRSPDDKLLASCADGTLKLWDATSANERKS 780
 Db 721 SDCFLKLDLWLNQECRNTMFGHTSNVHCRSPDDKLLASCADGTLKLWDATSANERKS 780
 Qy 781 INVQOFFLNLBEPQEDMEVIVKCCSWSDAGARIMVAANKIIFLFDIHTSGLLGEIHTGHH 840
 Db 781 INVQOFFLNLBEPQEDMEVIVKCCSWSDAGARIMVAANKIIFLFDIHTSGLLGEIHTGHH 840
 Qy 841 STIQYCDSPONHLAVVALSOYVELMNTDTSRQVADCRGHLVWGHVMPSPDGSSFLT 900
 Db 841 STIQYCDSPONHLAVVALSOYVELMNTDTSRQVADCRGHLVWGHVMPSPDGSSFLT 900
 Qy 901 SDDQTIRLWETKVCNKSAMVLMKQEDVVPQBNVWVLAVDHRRQLQINGRTQIDYLT 960
 Db 901 SDDQTIRLWETKVCNKSAMVLMKQEDVVPQBNVWVLAVDHRRQLQINGRTQIDYLT 960
 Qy 961 EAQVSCCLSPHLOVIAFGDENGAEIIELELVNRRIFQSRPQHKTVWHIQTADEKTLIS 1020
 Db 961 EAQVSCCLSPHLOVIAFGDENGAEIIELELVNRRIFQSRPQHKTVWHIQTADEKTLIS 1020
 Qy 1021 SSDDAEIQVWVWOLDKCIIFLGHQETVDFKFLKNSRLLSWSFDGTVKVMNIIITGNKEK 1080
 Db 1021 SSDDAEIQVWVWOLDKCIIFLGHQETVDFKFLKNSRLLSWSFDGTVKVMNIIITGNKEK 1080
 Qy 1081 FVCHQGTVLSCDISHDAKFSSTADAKTWSFDLPLPLHELGHNGCVCSAFSDVST 1140
 Db 1081 FVCHQGTVLSCDISHDAKFSSTADAKTWSFDLPLPLHELGHNGCVCSAFSDVST 1140
 Qy 1141 LLATGDDNGEIRIWNVNSGELLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1200
 Db 1141 LLATGDDNGEIRIWNVNSGELLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1200
 Qy 1201 WNVVVTGSSQTFYTNNGNLKKIHVSPDPKTYVTVDNLGILYILQTL 1248

Db 1201 WNVVVTGSSQTFYTNNGNLKKIHVSPDPKTYVTVDNLGILYILQTL 1248
 RESULT 7
 AAY97644
 ID AAY97644 standard; protein; 1248 AA.
 XX AAY97644;
 AC AAY97644;
 DT 20-APR-2001 (first entry)
 XX
 DE Apaf-1X-L-K63Q protein sequence.
 XX
 KW Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 KW hereditary disease; Apaf-1X-L-K63Q.
 XX
 OS Homo sapiens.
 XX
 FN WO200100827-A1.
 XX
 PD 04-JAN-2001.
 XX
 PP 30-JUN-2000; 2000WO-US018039.
 XX
 PR 30-JUN-1999; 99US-0141718P.
 XX
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Nunez G, Hu Y;
 XX
 DR WPI; 2001-112454/12.
 DR N-PSDB; AAA91123.
 XX
 PT Compositions for screening apoptosis pathway agonists and antagonists,
 PT useful for the treatment and diagnosis of cancer and autoimmune diseases,
 PT comprises new splice variants of wild-type Apaf-1 gene.
 XX
 PS Disclosure; Fig 33; 101pp; English.
 XX
 CC This sequence represents an Apaf-1 variant of the invention. The variants
 CC of the invention are all splice variants of a wild-type Apaf-1 gene (a
 CC mammalian homolog of CED-4 that participates in cytochrome c-dependent
 CC activation of caspase-3). The DNA is useful in screening assays for
 CC identifying apoptosis signalling pathway (antagonists, which are in turn
 CC useful as potential therapeutics and diagnostic or prognostic tools for
 CC diverse types of cancers, autoimmune diseases and hereditary diseases and
 CC for screening compounds that modulates the interaction of Apaf-1 with
 CC other members of the signalling pathway, i.e. their substrates or
 CC ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 CC cell-free assay system and the antibody generated to the translation
 CC product are used in immunoprecipitation assays to isolate new Apaf-1
 CC pathway constituents or their natural mutants
 XX
 SQ Sequence 1248 AA;
 Query Match 99.9%; Score 6615; DB 4; Length 1248;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1247; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MDKARNCLLOHRALEKDKITSYIMDMHISDGLTISEEKKVNEPTQQQRAAMLIKM 60
 Db 1 MDKARNCLLOHRALEKDKITSYIMDMHISDGLTISEEKKVNEPTQQQRAAMLIKM 60
 Qy 61 LKKQNDSSVSPYNALLHEGYKDLAALLHDGTPVVSSSGKDSVSGITSYVTVLCEGVP 120
 Db 61 LKKQNDSSVSPYNALLHEGYKDLAALLHDGTPVVSSSGKDSVSGITSYVTVLCEGVP 120
 Qy 121 QRPVVVTRKKLVNAIQKLSKLGEPGWVTHGAGCGKSVLAEEAVRDSHSLLEGCPG 180
 Db 121 QRPVVVTRKKLVNAIQKLSKLGEPGWVTHGAGCGKSVLAEEAVRDSHSLLEGCPG 180
 Qy 181 GVHWSVGKQKSGLLMKLQNLCTRLDQDSEFSQRLPLNIEBEAKDRILMLRKHPSLL 240

Db 181 GVHWVSQKDSGLMLKQLNCLRLDQDSFQRPLNIEAKDRILMLRKHPSLL 240
 Qy 241 ILDDVWDSWVLAQFQSQIILLTRDKSVTDSVMGPKYVVPVSSLGKGLKLSLFLVN 300
 Db 241 ILDDVWDSWVLAQFQSQIILLTRDKSVTDSVMGPKYVVPVSSLGKGLKLSLFLVN 300
 Qy 301 MKKADLPEQAHSIIKECKGSPVVSLLIGALLRFPNMEYLLQONKQKFRIRKSSVD 360
 Db 301 MKKADLPEQAHSIIKECKGSPVVSLLIGALLRFPNMEYLLQONKQKFRIRKSSVD 360
 Qy 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKQVVKVPTKVICILMDMETEEVEDILQEFVN 420
 Db 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKQVVKVPTKVICILMDMETEEVEDILQEFVN 420
 Qy 421 KSLFLCDNRGKSFYYLHDLQVDFPLETKNCSQQLDHLKKIITQFORVHOPHTLSPOEDC 480
 Db 421 KSLFLCDNRGKSFYYLHDLQVDFPLETKNCSQQLDHLKKIITQFORVHOPHTLSPOEDC 480
 Qy 481 MYWYNFLAYHMASAKHKLALMFSLDWKATKTELAVGPAHLIHEFVEYRHIIDEKDCAV 540
 Db 481 MYWYNFLAYHMASAKHKLALMFSLDWKATKTELAVGPAHLIHEFVEYRHIIDEKDCAV 540
 Qy 541 SENFQFLSLNGHILGROPENIVOLGCEPTESEVYQAKLOAKQEVDMGMLYLEWINK 600
 Db 541 SENFQFLSLNGHILGROPENIVOLGCEPTESEVYQAKLOAKQEVDMGMLYLEWINK 600
 Qy 601 KNTNLRLVVRPHDVAHYHACFSDGQRIASCADKTLQVFKATGEKLEIKAHEDV 660
 Db 601 KNTNLRLVVRPHDVAHYHACFSDGQRIASCADKTLQVFKATGEKLEIKAHEDV 660
 Qy 661 LCAFSTDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLATGS 720
 Db 661 LCAFSTDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLATGS 720
 Qy 721 SDCFLKMDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKMDATSANERKS 780
 Db 721 SDCFLKMDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKMDATSANERKS 780
 Qy 781 INVKQFFLNLEDPOEDMEVIVKCCSASDAGRIWAAKNKIFLFDIHTSGLLGEIHTGHH 840
 Db 781 INVKQFFLNLEDPOEDMEVIVKCCSASDAGRIWAAKNKIFLFDIHTSGLLGEIHTGHH 840
 Qy 841 STIQYCDFSPQNLAVVALSQYCVELWNTDSRKVADCRGHLVWVGFSPDGSFLTS 900
 Db 841 STIQYCDFSPQNLAVVALSQYCVELWNTDSRKVADCRGHLVWVGFSPDGSFLTS 900
 Qy 901 SDDQIRLWETKVKCKNSAVMLKQEVVDVVFQENWVWVLAVDHRRLOLINGRTGQIDYLT 960
 Db 901 SDDQIRLWETKVKCKNSAVMLKQEVVDVVFQENWVWVLAVDHRRLOLINGRTGQIDYLT 960
 Qy 961 EAQVSCCLSPHLOVIAFGDENGAIETLELVNRRIFQSRFOHKTVMHIOFTADEKTLIS 1020
 Db 961 EAQVSCCLSPHLOVIAFGDENGAIETLELVNRRIFQSRFOHKTVMHIOFTADEKTLIS 1020
 Qy 1021 SDDDAEIQVWAKWLDKCIFLRGHQTVDKFRLLKNSRLLSSPDGTQVKNVNIITGNKEK 1080
 Db 1021 SDDDAEIQVWAKWLDKCIFLRGHQTVDKFRLLKNSRLLSSPDGTQVKNVNIITGNKEK 1080
 Qy 1081 FVCHQGTVLSCDISHDATKFSSTSAKTAKIWSFDLLPLPLHELGHNGCVRCAPSVDS 1140
 Db 1081 FVCHQGTVLSCDISHDATKFSSTSAKTAKIWSFDLLPLPLHELGHNGCVRCAPSVDS 1140
 Qy 1141 LIATGDDNGEIRIWNVNGELLHLCAPLSEGAATHGGVNTDLCFSPDGKXLSAGGYIK 1200
 Db 1141 LIATGDDNGEIRIWNVNGELLHLCAPLSEGAATHGGVNTDLCFSPDGKXLSAGGYIK 1200
 Qy 1201 WNVVVTGESSQFTYNTGNTLKKIHVSDFPKTYVVDNLGILYILQTLLE 1248
 Db 1201 WNVVVTGESSQFTYNTGNTLKKIHVSDFPKTYVVDNLGILYILQTLLE 1248

RESULT 8

AAAY97645
 ID AAY97645 standard; protein; 1248 AA.
 XX
 AC AAY97645;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Apaf-1XL-L83A protein sequence.
 XX
 KW Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 KW hereditary disease; Apaf-1XL-L83A.
 XX
 OS Homo sapiens.
 XX
 PN WO200100827-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 30-JUN-2000; 2000WO-US018039.
 XX
 PR 30-JUN-1999; 99US-0141718P.
 XX
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Nunez G, Hu Y;
 XX
 PP WPI; 2001-112454/12.
 XX
 DR N-PSDB; AAA91124.
 XX
 PT Compositions for screening apoptosis pathway agonists and antagonists,
 PT useful for the treatment and diagnosis of cancer and autoimmune diseases,
 PT comprises new splice variants of wild-type Apaf-1 gene.
 XX
 PS Disclosure; Fig 34; 101pp; English.
 XX
 CC This sequence represents an Apaf-1 variant of the invention. The variants
 CC of the invention are all splice variants of a wild-type Apaf-1 gene (a
 CC mammalian homolog of CED-4 that participates in cytochrome c-dependent
 CC activation of caspase-3). The DNA is useful in screening assays for
 CC identifying apoptosis signalling pathway (antagonists, which are in turn
 CC useful as potential therapeutics and diagnostic or prognostic tools for
 CC diverse types of cancers, autoimmune diseases and hereditary diseases and
 CC for screening compounds that modulate the interaction of Apaf-1 with
 CC other members of the signalling pathway, i.e. their substrates or
 CC ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 CC cell-free assay system and the antibody generated to the translation
 CC product are used in immunoprecipitation assays to isolate new Apaf-1
 CC pathway constituents or their natural mutants
 XX
 SQ Sequence 1248 AA;
 Query Match 99.9%; Score 6614; DB 4; Length 1248;
 Best Local Similarity 99.9%; Fred. No. 0;
 Matches 1247; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MDKARNCILQHREALEKDIKTSYIMDMISDGLFTTISEEKVNEPTQOQRAALIM 60
 Db 1 MDKARNCILQHREALEKDIKTSYIMDMISDGLFTTISEEKVNEPTQOQRAALIM 60
 Qy 61 LKQNDYSVSYFNALLHEGYKDLAALHDGIPVVSSSGKDSVSGITSYRTVLCGGVP 120
 Db 61 LKQNDYSVSYFNALLHEGYKDLAALHDGIPVVSSSGKDSVSGITSYRTVLCGGVP 120
 Qy 121 QRPVVFVTRKLVNAIQKLSKNGCPQWVTIHGMACGKSVLAABAVRDHSLLEGCPFG 180
 Db 121 QRPVVFVTRKLVNAIQKLSKNGCPQWVTIHGMACGKSVLAABAVRDHSLLEGCPFG 180
 Qy 181 GVHWVSQKDSGLMLKQLNCLRLDQDSFQRPLNIEAKDRILMLRKHPSLL 240
 Db 181 GVHWVSQKDSGLMLKQLNCLRLDQDSFQRPLNIEAKDRILMLRKHPSLL 240
 Qy 241 ILDDVWDSWVLAQFQSQIILLTRDKSVTDSVMGPKYVVPVSSLGKGLKLSLFLVN 300

Db 241 ILDDVWDSWLVKAFDSQCIILLTRDKSVTVSVGMGPKVVPVSSSLGKEKLEILSLFVN 300
 Qy 301 MKKADLPQAHSIIEKEKSGPLVSLIGALLRDFENRWEYLLKQLNQKQKRIKSSSYD 360
 Db 301 MKKADLPQAHSIIEKEKSGPLVSLIGALLRDFENRWEYLLKQLNQKQKRIKSSSYD 360
 Qy 361 YEALDEAMSIIVEMLRDIKDYITDLSILQKDVKPTKVLCLMDMETEVEDILQEFVN 420
 Db 361 YEALDEAMSIIVEMLRDIKDYITDLSILQKDVKPTKVLCLMDMETEVEDILQEFVN 420
 Qy 421 KSLFCDNRGKSFRIYLDHLDVDFTEKNCQLODLHKKIITQFORYHOPHTLSPDQDC 480
 Db 421 KSLFCDNRGKSFRIYLDHLDVDFTEKNCQLODLHKKIITQFORYHOPHTLSPDQDC 480
 Qy 481 MYWYNFLAYHMASAKWHKELCALMFLDWIKAKTELVGPAHLIHEFVEYRHLDEKCAV 540
 Db 481 MYWYNFLAYHMASAKWHKELCALMFLDWIKAKTELVGPAHLIHEFVEYRHLDEKCAV 540
 Qy 541 SENTQEFSLNGHLLGROPPFNIIVGLGCEPETSVEYQOAKLOAKQEVNDGMLYLEWINK 600
 Db 541 SENTQEFSLNGHLLGROPPFNIIVGLGCEPETSVEYQOAKLOAKQEVNDGMLYLEWINK 600
 Qy 601 KNTNLSRLVVRPHTDAVYHACFSEDGRIASCGADKTLQVFKATGKLEIKAHDEV 660
 Db 601 KNTNLSRLVVRPHTDAVYHACFSEDGRIASCGADKTLQVFKATGKLEIKAHDEV 660
 Qy 661 LCCAFSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEVQNCCHFTNSHHLLATGS 720
 Db 661 LCCAFSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEVQNCCHFTNSHHLLATGS 720
 Qy 721 SDCELKLDLWLNQKCRNTMFGHTSNVHCRSPDDKLLASCSADGTLKLDATSANERKS 780
 Db 721 SDCELKLDLWLNQKCRNTMFGHTSNVHCRSPDDKLLASCSADGTLKLDATSANERKS 780
 Qy 781 INVQKFFLNLRDPOEDMEVIVKCCSWASADGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 840
 Db 781 INVQKFFLNLRDPOEDMEVIVKCCSWASADGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 840
 Qy 841 STIQCDFSPQNHVALVVALSQCVELMNTDSRSKVDRCRHLVWGVWSPDGSSFLTS 900
 Db 841 STIQCDFSPQNHVALVVALSQCVELMNTDSRSKVDRCRHLVWGVWSPDGSSFLTS 900
 Qy 901 SDDQIRLWETKVCXKSAVLMKQEVVVFQENEVMLAVDHIRRLQLINGRTQIDYLT 960
 Db 901 SDDQIRLWETKVCXKSAVLMKQEVVVFQENEVMLAVDHIRRLQLINGRTQIDYLT 960
 Qy 961 EAOVSCCLSPHLOVIAFGDENGAIETILELVNNRIFQSRFOHKKTVWHIQTADKTLIS 1020
 Db 961 EAOVSCCLSPHLOVIAFGDENGAIETILELVNNRIFQSRFOHKKTVWHIQTADKTLIS 1020
 Qy 1021 SSDDAEIQVNNWQDKCIFLFGHQBTVKDFRLLKNSRLLSWSFDGTVKVWNIITGNKEKD 1080
 Db 1021 SSDDAEIQVNNWQDKCIFLFGHQBTVKDFRLLKNSRLLSWSFDGTVKVWNIITGNKEKD 1080
 Qy 1081 FVCHQGTLSCDISHDATKFSSTADKTAKTWSFDLLPLHELGHNGCVRCASFVDS 1140
 Db 1081 FVCHQGTLSCDISHDATKFSSTADKTAKTWSFDLLPLHELGHNGCVRCASFVDS 1140
 Qy 1141 LLATGDDNGEIRIWNVSGELLHLICAPLSEGAATHGCGWTDLCFSPDGKMLISAGGYIK 1200
 Db 1141 LLATGDDNGEIRIWNVSGELLHLICAPLSEGAATHGCGWTDLCFSPDGKMLISAGGYIK 1200
 Qy 1201 WNNVVTGESSOTFTYNGTNLKKIHVSDFKTYVTVDNLGILYILOTLE 1248
 Db 1201 WNNVVTGESSOTFTYNGTNLKKIHVSDFKTYVTVDNLGILYILOTLE 1248

RESULT 9
 ID AA97641
 standard; protein; 1248 AA.
 AC AA97641;

XX 20-APR-2001 (first entry)
 XX Apaf-1XL-L10A protein sequence.
 XX Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 KW hereditary disease; Apaf-1XL-L10A.
 XX Homo sapiens.
 XX WO200100827-A1.
 XX 04-JAN-2001.
 XX 30-JUN-2000; 2000WO-US018039.
 XX 30-JUN-1999; 99US-0141718P.
 XX (UNMI) UNIV MICHIGAN.
 PA Nunez G, Hu Y;
 XX WPI; 2001-112454/12.
 XX N-PSDB; AAA91120.
 CC Compositions for screening apoptosis pathway agonists and antagonists,
 CC useful for the treatment and diagnosis of cancer and autoimmune diseases,
 CC comprises new splice variants of wild-type Apaf-1 gene.
 PS Disclosure; Fig 30; 101pp; English.
 CC This sequence represents an Apaf-1 variant of the invention. The variants
 CC of the invention are all splice variants of a wild-type Apaf-1 gene (a
 CC mammalian homolog of CED-4 that participates in cytochrome c-dependent
 CC activation of caspase-3). The DNA is useful in screening assays for
 CC identifying apoptosis signalling pathway (antagonists, which are in turn
 CC useful as potential therapeutics and diagnostic or prognostic tools for
 CC diverse types of cancers, autoimmune diseases and hereditary diseases and
 CC for screening compounds that modulates the interaction of Apaf-1 with
 CC other members of the signalling pathway, i.e. their substrates or
 CC ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 CC cell-free assay system and the antibody generated to the translation
 CC product are used in immunoprecipitation assays to isolate new Apaf-1
 CC pathway constituents or their natural mutants
 XX Sequence 1248 AA;

Query Match 99.9%; Score 6614; DB 4; Length 1248;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1247; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MDKARVCLQNRALKEDIKTSYIMDHMSDGLTISEEEKVNRNEPTQOORAAALIKMI 60
 Db 1 MDKARNVCLQNRALKEDIKTSYIMDHMSDGLTISEEEKVNRNEPTQOORAAALIKMI 60
 Qy 61 LKXNDSDSVSYFNALLHEGYKDLAALLHDGIPVYSSSSGKDSVSGITSVYRTVLCGGVP 120
 Db 61 LKXNDSDSVSYFNALLHEGYKDLAALLHDGIPVYSSSSGKDSVSGITSVYRTVLCGGVP 120
 Qy 121 QRPVVFTRKKLVNAIQKLSKLGEPGWVTHGMAGCGKSVLAABAVRDSHSLLEGCPGP 180
 Db 121 QRPVVFTRKKLVNAIQKLSKLGEPGWVTHGMAGCGKSVLAABAVRDSHSLLEGCPGP 180
 Qy 181 GVHWVSVGKODKSGLLMKLQNLCTRLDQDSFSORLPNTEEAKDRILMLRKHPSLL 240
 Db 181 GVHWVSVGKODKSGLLMKLQNLCTRLDQDSFSORLPNTEEAKDRILMLRKHPSLL 240
 Qy 241 ILDDVWDSWLVKAFDSQCIILLTRDKSVTVSVGMGPKVVPVSSSLGKEKLEILSLFVN 300
 Db 241 ILDDVWDSWLVKAFDSQCIILLTRDKSVTVSVGMGPKVVPVSSSLGKEKLEILSLFVN 300
 Qy 301 MKKADLPQAHSIIEKEKSGPLVSLIGALLRDFENRWEYLLKQLNQKQKRIKSSSYD 360

Db 301 MKKADLPEQAHSIIKECKGSLVLSLIGALLRFPNRYEYLKQONQKFRIRKSSSYD 360
 Qy 361 YEALDEAMSISVEMLRERIDKYDYLTDLSILOQKQVKTPLVLCILMDMETEEVEDILOEFVN 420
 Db 361 YEALDEAMSISVEMLRERIDKYDYLTDLSILOQKQVKTPLVLCILMDMETEEVEDILOEFVN 420
 Qy 421 KSLFLCDNRNGKFRYYLHDLQVDFLTCKNGSQDLHKKIITQFORVHQPHTLSPDOEDC 480
 Db 421 KSLFLCDNRNGKFRYYLHDLQVDFLTCKNGSQDLHKKIITQFORVHQPHTLSPDOEDC 480
 Qy 481 MYWYNFLAYHMASAKWHELCALMFSLDWKATKELVGPFAHLIHEFVEYRHILDEKCAV 540
 Db 481 MYWYNFLAYHMASAKWHELCALMFSLDWKATKELVGPFAHLIHEFVEYRHILDEKCAV 540
 Qy 541 SENFOEFLSLNGHLGRQFFNIYVQLGCEPTESEVYQAKLOAQKQEVNDGMLYEWINK 600
 Db 541 SENFOEFLSLNGHLGRQFFNIYVQLGCEPTESEVYQAKLOAQKQEVNDGMLYEWINK 600
 Qy 601 KNITNLRLVVRPHTDAVYHACFSEDOGTASCGADKTLQVFKATGKLEIKKAEHEDEV 660
 Db 601 KNITNLRLVVRPHTDAVYHACFSEDOGTASCGADKTLQVFKATGKLEIKKAEHEDEV 660
 Qy 661 LCCAFSTDRFATCSVDKKVKNWMTGELVYTDHSEQVNCCHFTNSHHLLATGS 720
 Db 661 LCCAFSTDRFATCSVDKKVKNWMTGELVYTDHSEQVNCCHFTNSHHLLATGS 720
 Qy 721 SDCLFLKDLNOKECRNTMFGHTNSVNHCRFSPDDKLLASCSADGTLKMDATSANERKS 780
 Db 721 SDCLFLKDLNOKECRNTMFGHTNSVNHCRFSPDDKLLASCSADGTLKMDATSANERKS 780
 Qy 781 INVQKPFLEDPQEDMEVIVKCCMSADGARIMAAKNKIFLDIHTSGLLGEIHTGHH 840
 Db 781 INVQKPFLEDPQEDMEVIVKCCMSADGARIMAAKNKIFLDIHTSGLLGEIHTGHH 840
 Qy 841 STIQYCDSPQNHLLAVALSQYCVELWNTDSRSKVADCRGHLVSWHGMFSPDGSFLTS 900
 Db 841 STIQYCDSPQNHLLAVALSQYCVELWNTDSRSKVADCRGHLVSWHGMFSPDGSFLTS 900
 Qy 901 SDQOTIRLWETKVKCKNSAVMLKQEVNDVVFQENEVMVLAVDHIRRLQIINGRTQIDYLT 960
 Db 901 SDQOTIRLWETKVKCKNSAVMLKQEVNDVVFQENEVMVLAVDHIRRLQIINGRTQIDYLT 960
 Qy 961 EAQVSCCLSLPHLOVIAFGDENGAIETLELVNRRIFQSRFQHKKTVMHIOFTADEKTLIS 1020
 Db 961 EAQVSCCLSLPHLOVIAFGDENGAIETLELVNRRIFQSRFQHKKTVMHIOFTADEKTLIS 1020
 Qy 1021 SSDDAEIQVWNLDCIKFLRGHOETVKDFLLKNSRLLSWSFDGTWKVWNIIITGNKED 1080
 Db 1021 SSDDAEIQVWNLDCIKFLRGHOETVKDFLLKNSRLLSWSFDGTWKVWNIIITGNKED 1080
 Qy 1081 FVCHQGTVLSCDI SHDATKFSSTADKATKINSFDLILPLHELCHNGCVRCASFVDS 1140
 Db 1081 FVCHQGTVLSCDI SHDATKFSSTADKATKINSFDLILPLHELCHNGCVRCASFVDS 1140
 Qy 1141 LLATGDDNGEIRINWVNSGELLHLCAPLSEGAATHGQVWTDLCFSPDGMKMLISAGGYIK 1200
 Db 1141 LLATGDDNGEIRINWVNSGELLHLCAPLSEGAATHGQVWTDLCFSPDGMKMLISAGGYIK 1200
 Qy 1201 WNNVVTGESSQFTYNGTNLKKIHVSPDKFTYVVDNLGILYILOTLE 1248
 Db 1201 WNNVVTGESSQFTYNGTNLKKIHVSPDKFTYVVDNLGILYILOTLE 1248

RESULT 10
 ID AAY97647
 AC AAY97647
 XX AAY97647;
 XX 20-APR-2001 (first entry)
 DT
 XX
 DE Apaf-1XL-M368L/K160R protein sequence.

XX Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 KW hereditary disease; Apaf-1XL-M368L/K160R.
 OS Homo sapiens.
 XX WC200100827-A1.
 XX 04-JAN-2001.
 XX 30-JUN-2000; 2000WO-US018039.
 XX 30-JUN-1999; 99US-0141718P.
 XX (UNMI) UNIV MICHIGAN.
 XX Nunez G, Hu Y;
 XX WPI: 2001-112454/12.
 XX N-PSDB; AAN91126.
 XX Compositions for screening apoptosis pathway agonists and antagonists,
 useful for the treatment and diagnosis of cancer and autoimmune diseases,
 comprises new splice variants of wild-type Apaf-1 gene.
 PS Disclosure; Fig 36; 101pp; English.
 XX This sequence represents an Apaf-1 variant of the invention. The variants
 of the invention are all splice variants of a wild-type Apaf-1 gene (a
 mammalian homolog of CED-4 that participates in cytochrome c-dependent
 activation of caspase-3). The DNA is useful in screening assays for
 identifying apoptosis signalling pathway (antagonists, which are in turn
 useful as potential therapeutic and diagnostic or prognostic tools for
 diverse types of cancers, autoimmune diseases and hereditary diseases and
 for screening compounds that modulates the interaction of Apaf-1 with
 other members of the signalling pathway, i.e. their substrates or
 ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 cell-free assay system and the antibody generated to the translation
 product are used in immunoprecipitation assays to isolate new Apaf-1
 pathway constituents or their natural mutants

Query Match 99.9%; Score 6613; DB 4; Length 1248;
 Best Local Similarity 99.8%; Pred No. 0;
 Matches 1246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MDKARNCLLOHREALEKDKITSYIMDHWSIDGFLTISEEKVNEPTQOQRAAMLIKMI 60
 Db 1 MDKARNCLLOHREALEKDKITSYIMDHWSIDGFLTISEEKVNEPTQOQRAAMLIKMI 60
 Qy 61 LKKNDSYVSYFYNALLHREGYKOLAAALHDGIPVYSSSGKDSVSGITSYVTVLCEGVP 120
 Db 61 LKKNDSYVSYFYNALLHREGYKOLAAALHDGIPVYSSSGKDSVSGITSYVTVLCEGVP 120
 Qy 121 QRPVVFTRKKLVNAIQOKSLKKGEPQVTHGMAGGKSVLAABAVRDHSLSGCPFG 180
 Db 121 QRPVVFTRKKLVNAIQOKSLKKGEPQVTHGMAGGKSVLAABAVRDHSLSGCPFG 180
 Qy 181 GVHWVSVGKQDSGLLMKLNQLCTRLDQDESFSQRLPLNIEEAKDLRLILMLRKHPRSL 240
 Db 181 GVHWVSVGKQDSGLLMKLNQLCTRLDQDESFSQRLPLNIEEAKDLRLILMLRKHPRSL 240
 Qy 241 ILDDVDSWLVKAFDSQCCQLITTRDKSVTDSVNGPKYVYVVSSESLGKGLIILSFVN 300
 Db 241 ILDDVDSWLVKAFDSQCCQLITTRDKSVTDSVNGPKYVYVVSSESLGKGLIILSFVN 300
 Qy 301 MKKADLPEQAHSIIKECKGSLVLSLIGALLRFPNRYEYLKQONQKFRIRKSSSYD 360
 Db 301 MKKADLPEQAHSIIKECKGSLVLSLIGALLRFPNRYEYLKQONQKFRIRKSSSYD 360
 Qy 361 YEALDEAMSISVEMLRERIDKYDYLTDLSILOQKQVKTPLVLCILMDMETEEVEDILOEFVN 420

Db 361 YEALDEALSIIVEMLRDIDKDYITDLSILQKDVKPTKVLCLMDMETEEVEDILOQFVN 420
 Qy 421 KSLFCDNRNGKSFYLLHDLQVDFLTERKNSQLQDLHKKIITQFQYHQPHTLSPDQEDC 480
 Db 421 KSLFCDNRNGKSFYLLHDLQVDFLTERKNSQLQDLHKKIITQFQYHQPHTLSPDQEDC 480
 Qy 481 MYTNFLAYHNASAKMKELCALMPSLDWIIKAKTELGPAPLIIHEFVEYRHILDEKCAV 540
 Db 481 MYTNFLAYHNASAKMKELCALMPSLDWIIKAKTELGPAPLIIHEFVEYRHILDEKCAV 540
 Qy 541 SENFOELSLINGHLIGRPPNNIIVOLGICEPETSVEVQOAKLOAKQEVNDGMLYLEWINK 600
 Db 541 SENFOELSLINGHLIGRPPNNIIVOLGICEPETSVEVQOAKLOAKQEVNDGMLYLEWINK 600
 Qy 601 KNIITNLSRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFKAGTCKLEIKAHEDV 660
 Db 601 KNIITNLSRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFKAGTCKLEIKAHEDV 660
 Qy 661 LCCAFSTDDRIATCSVDKVKIWNMTGELVHYTDEHSEOVNCHFTNSHHLLATGS 720
 Db 661 LCCAFSTDDRIATCSVDKVKIWNMTGELVHYTDEHSEOVNCHFTNSHHLLATGS 720
 Qy 721 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
 Db 721 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
 Qy 781 INVQFFLNLEDPQEDMEVIVKCCSWASADGARIWVAANKIIFLFDIHTSGLLGHIHTGH 840
 Db 781 INVQFFLNLEDPQEDMEVIVKCCSWASADGARIWVAANKIIFLFDIHTSGLLGHIHTGH 840
 Qy 841 STIQYCFSPONHLAVVALSYCYVELMNTDSRSKVADCRHLSVWGVMSFPGSSFLTS 900
 Db 841 STIQYCFSPONHLAVVALSYCYVELMNTDSRSKVADCRHLSVWGVMSFPGSSFLTS 900
 Qy 901 SDDQITRLWETKVKCKNSAVMLKQEVVVOENVMVLAVDHIIRLQILNGRTQIDYLT 960
 Db 901 SDDQITRLWETKVKCKNSAVMLKQEVVVOENVMVLAVDHIIRLQILNGRTQIDYLT 960
 Qy 961 EAQVSCCLSPHLOVIAFGDENGAIETILELVNNRIFQSRFOHKKTVMHIOFTADEKTLIS 1020
 Db 961 EAQVSCCLSPHLOVIAFGDENGAIETILELVNNRIFQSRFOHKKTVMHIOFTADEKTLIS 1020
 Qy 1021 SSDDAEIQVNWQDKICIFLGHQETVKDFLLKNSRLLSWSPDGTVKVWNIITGNKEK 1080
 Db 1021 SSDDAEIQVNWQDKICIFLGHQETVKDFLLKNSRLLSWSPDGTVKVWNIITGNKEK 1080
 Qy 1081 FVCHOGTVLSCDISHDATKFSSTADKTAKTWSFDLILPLHELGHNGCVRCASFVDST 1140
 Db 1081 FVCHOGTVLSCDISHDATKFSSTADKTAKTWSFDLILPLHELGHNGCVRCASFVDST 1140
 Qy 1141 LLATGDDNGEIRIWNVSGELLHLICAPLSEBGAATHGCVWTDLCFSPDGKMLISAGGYIK 1200
 Db 1141 LLATGDDNGEIRIWNVSGELLHLICAPLSEBGAATHGCVWTDLCFSPDGKMLISAGGYIK 1200
 Qy 1201 WNVVTGESSQTFYTNGLNKKIHSVDPFKTYTVDNLGILYIILQTL 1248
 Db 1201 WNVVTGESSQTFYTNGLNKKIHSVDPFKTYTVDNLGILYIILQTL 1248

RESULT 11
 ID AAY97642
 AC AAY97642 standard; protein; 1248 AA.
 XX AAY97642;
 XX 20-APR-2001 (first entry)
 XX Apaf-1X-D27A protein sequence.
 XX Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 KW hereditary disease; Apaf-1X-D27A.

XX Homo sapiens.
 OS WO200100827-A1.
 XX 04-JAN-2001.
 XX 30-JUN-2000; 2000WO-US018039.
 XX 30-JUN-1999; 99US-0141718P.
 XX (UNWI) UNIV MICHIGAN.
 XX Nunez G, Hu Y;
 XX WPI; 2001-112454/12.
 XX N-PSDB; AAA91121.
 XX Compositions for screening apoptosis pathway agonists and antagonists,
 useful for the treatment and diagnosis of cancer and autoimmune diseases,
 comprises new splice variants of wild-type Apaf-1 gene.
 XX Disclosure; Fig 31; 101pp; English.
 XX This sequence represents an Apaf-1 variant of the invention. The variants
 of the invention are all splice variants of a wild-type Apaf-1 gene (a
 mammalian homolog of CED-4 that participates in cytochrome c-dependent
 activation of caspase-3). The DNA is useful in screening assays for
 identifying apoptosis signalling pathway (antagonists, which are in turn
 useful as potential therapeutics and diagnostic or prognostic tools for
 diverse types of cancers, autoimmune diseases and hereditary diseases and
 for screening compounds that modulates the interaction of Apaf-1 with
 other members of the signalling pathway, i.e. their substrates or
 ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 cell-free assay system and the antibody generated to the translation
 product are used in immunoprecipitation assays to isolate new Apaf-1
 pathway constituents or their natural mutants
 XX Sequence 1248 AA;

Query Match 99.9%; Score 6611; DB 4; Length 1248;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 124; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 MDKARNCLLQHREALEKDIKTSYIMDMHISDGLTISEEEKVNEPTQQQRAAMLKMI 60
 Db 1 MDKARNCLLQHREALEKDIKTSYIMDMHISDGLTISEEEKVNEPTQQQRAAMLKMI 60
 Qy 61 LKQNDYSYFYNALLHEGYKDLAALLHDGIPVVSSSSGKDSVSGITSYVTVLCEGGVP 120
 Db 61 LKQNDYSYFYNALLHEGYKDLAALLHDGIPVVSSSSGKDSVSGITSYVTVLCEGGVP 120
 Qy 121 QRPVVFVTRKKLVNAIOOKLSKLGEPGWTHGMAGCGSVLAEEAVRDSHSLGCPG 180
 Db 121 QRPVVFVTRKKLVNAIOOKLSKLGEPGWTHGMAGCGSVLAEEAVRDSHSLGCPG 180
 Qy 181 GVHWVSYGKQDKSGLLMKQLNCTRLQDQSFQSLPLNIEEAKDRILMLRKHPSLL 240
 Db 181 GVHWVSYGKQDKSGLLMKQLNCTRLQDQSFQSLPLNIEEAKDRILMLRKHPSLL 240
 Qy 241 ILDDVDSWVWKAPDSQCQILLTTRDKSVTSDVNGPKYVVPVSSSLGKGLTSLFVN 300
 Db 241 ILDDVDSWVWKAPDSQCQILLTTRDKSVTSDVNGPKYVVPVSSSLGKGLTSLFVN 300
 Qy 301 MKKADLPEQAHSIIECKCGSPLVVLIGALLRDPFNWYVLLKOLQNKQPKIRKSSSYD 360
 Db 301 MKKADLPEQAHSIIECKCGSPLVVLIGALLRDPFNWYVLLKOLQNKQPKIRKSSSYD 360
 Qy 361 YEALDEAMSIIVEMLRDIDKDYITDLSILQKDVKPTKVLCLMDMETEEVEDILOQFVN 420
 Db 361 YEALDEAMSIIVEMLRDIDKDYITDLSILQKDVKPTKVLCLMDMETEEVEDILOQFVN 420
 Qy 421 KSLFCDNRNGKSFYLLHDLQVDFLTERKNSQLQDLHKKIITQFQYHQPHTLSPDQEDC 480

Db 470 MYWNLFLAYHNASAKHKLALCALPESLDWIIKAKTELVGPAHLIHEFYRHLDEKCAV 529
Qy 541 SENFQEFSLNGLHGRPFNNIVOLGLCEPETSEYQOAKLOAEVDNGLYLEWINK 600
Db 530 SENFQEFSLNGLHGRPFNNIVOLGLCEPETSEYQOAKLOAEVDNGLYLEWINK 589
Qy 601 KNIITNLSLVVRPHTDAVYHACFSDGQRIASCADKTLQVFKAPTGBKLLEIKAHEDV 660
Db 590 KNIITNLSLVVRPHTDAVYHACFSDGQRIASCADKTLQVFKAPTGBKLLEIKAHEDV 649
Qy 661 LCCAFSTDDRFATCSVDKVKIWNMTGELVHTYDEHSEVNCCHFTNSHHLLATGS 720
Db 650 LCCAFSTDDRFATCSVDKVKIWNMTGELVHTYDEHSEVNCCHFTNSHHLLATGS 709
Qy 721 SDCFLKLDLWLNQKECRNTMFGHTSNVHCRFSPDDKLLASCADTKLWDATSANERKS 780
Db 710 SDCFLKLDLWLNQKECRNTMFGHTSNVHCRFSPDDKLLASCADTKLWDATSANERKS 769
Qy 781 INVQOFFNLNLEDPQEDMEVIVKCCSWASADGARIMVAANKIPLFDIHTSGLLGEIHTGHH 840
Db 770 INVQOFFNLNLEDPQEDMEVIVKCCSWASADGARIMVAANKIPLFDIHTSGLLGEIHTGHH 829
Qy 841 STIOYCDSPONHLAVVALSOYVELWNTDERSKVADCRGHLVWGVNFPDGSSEFLTS 900
Db 830 STIOYCDSPONHLAVVALSOYVELWNTDERSKVADCRGHLVWGVNFPDGSSEFLTS 889
Qy 901 SDDQIRLWETKVKCKSNVLMKQEDVVVFOENEVWVLAVDHRRQLQINGRTGQIDYLT 960
Db 890 SDDQIRLWETKVKCKSNVLMKQEDVVVFOENEVWVLAVDHRRQLQINGRTGQIDYLT 949
Qy 961 EAQVSCCLSPHLOVIAFGDENGAEILELVNRRIFQSRFQKKTVMHIOFTADEKTLIS 1020
Db 950 EAQVSCCLSPHLOVIAFGDENGAEILELVNRRIFQSRFQKKTVMHIOFTADEKTLIS 1009
Qy 1021 SSDAEIQVWVWQDKCIFLGRHOETVKDFLLKXNRLLSWSPDGTVKVWNIITGNKEKD 1080
Db 1010 SSDAEIQVWVWQDKCIFLGRHOETVKDFLLKXNRLLSWSPDGTVKVWNIITGNKEKD 1069
Qy 1081 FVCHQGTVLSGDISHDATKFSSTADKTAKTWSFDLLPLHELHGHNGCVCSAFSDVST 1140
Db 1070 FVCHQGTVLSGDISHDATKFSSTADKTAKTWSFDLLPLHELHGHNGCVCSAFSDVST 1129
Qy 1141 LLATGDDNGEIRIWNVNSGELLHLICAPLSEGAATHGQWVTDLCFSPDGRMLISAGGYIK 1200
Db 1130 LLATGDDNGEIRIWNVNSGELLHLICAPLSEGAATHGQWVTDLCFSPDGRMLISAGGYIK 1189
Qy 1201 WNVVVTGESSQTFYTNGLTKKIHVSPDFKTYVTVDNLGILYILQTL 1248
Db 1190 WNVVVTGESSQTFYTNGLTKKIHVSPDFKTYVTVDNLGILYILQTL 1237
RESULT 13
ADP65308
ID ADP65308 standard; protein; 1237 AA.
XX AC ADP65308;
XX AC ADP65308;
DT 12-AUG-2004 (first entry)
XX Human apoptotic protease activating factor isoform A, apoptotic protease.
DE autoimmune disease; arthritis; gene expression analysis;
KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW immune; human.
OS Homo sapiens.
XX
FN W02003072827-A1.
XX

PD 04-SEP-2003.
XX 31-OCT-2002; 2002WO-US035433.
PF 31-OCT-2001; 2001US-0336220P.
PR (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
PA Hirsch R, Thorton SL;
XX WPI; 2003-712740/67.
PI GENBANK; NP_037361.
XX
DR Diagnosing and analyzing autoimmune disease using gene expression
XX profiles and microarray technology, useful for diagnosing and treating
PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
PT gout.
XX
XX Disclosure; Page; 56pp; English.
XX
XX The invention relates to a novel method for diagnosing and analysing
CC autoimmune disease or arthritides. The method comprises obtaining a
CC patient sample containing mRNA, analysing gene expression using the mRNA
CC that results in a gene expression signature of the mRNA, and using that
CC gene expression signature to diagnose or analyse the autoimmune disease
CC or arthritides in the patient, where gene expression of at least 60% of
CC the genes correlates with that of the gene signature. The invention
CC further comprises: a treatment of rheumatoid arthritis; identification of
CC genes for targeting in the treatment of rheumatoid arthritis in a mammal
CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
CC array or gene chip, specific for rheumatoid arthritis; diagnosis or
CC analyses of autoimmune disease or rheumatoid arthritis; screening the
CC efficacy of a candidate drug in vitro for the treatment of collagen-
CC induced arthritis; and reducing the symptoms associated with collagen-
CC induced arthritis. The compositions of the invention have the following
CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
CC antigout, antiinflammatory, dermatological, and immunomodulatory. The
CC methods and compositions of the present invention are useful for
CC diagnosing and treating autoimmune disease or arthritides, such as
CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
CC immune disease caused by an infectious agent. This sequence represents a
CC protein sequence relating to the genes used in the analysis and treatment
CC of autoimmune diseases or arthritides. Note: This sequence is not shown
CC in the specification. It has been supplied in an electronic format from
CC WIPO.
XX
SQ Sequence 1237 AA;
Query Match 98.9%; Score 6548.5; DB 7; Length 1237;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1237; Conservative 0; Mismatches 0; Indels 11; Gaps 1;
Qy 1 MDAKARNCLLOHREALEKDIKTSYIMDMHISDGLFTTSEEKVNREPTQOQRAAMLIKI 60
Db 1 MDAKARNCLLOHREALEKDIKTSYIMDMHISDGLFTTSEEKVNREPTQOQRAAMLIKI 60
Qy 61 LKDNDSVSVFYNALHHEGYKDLAALLHDGIPVYSSSSGKDSVSGITSYVTVLCEGGVP 120
Db 61 LKDNDSVSVFYNALHHEGYKDLAALLHDGIPVYSSSSGKDSVSGITSYVTVLCEGGVP 109
Qy 121 QRPVVFVTRKLVNVAIQKLSKLBEGFGWVTHGMAGCKSVLAEEAVRDHSLLEGCPFG 180
Db 110 QRPVVFVTRKLVNVAIQKLSKLBEGFGWVTHGMAGCKSVLAEEAVRDHSLLEGCPFG 169
Qy 181 GVHWSVGKQDKSGLLMKLQNLCTRLQDQDSFSQRLPLNTEAKDRILMLRKHPRSL 240
Db 170 GVHWSVGKQDKSGLLMKLQNLCTRLQDQDSFSQRLPLNTEAKDRILMLRKHPRSL 229
Qy 241 ILDDVWDSWLVKAFDSDQCIILLTTRDKSVTDSVMGPKYVVPVSSLGKGLTSLFPVN 300
Db 230 ILDDVWDSWLVKAFDSDQCIILLTTRDKSVTDSVMGPKYVVPVSSLGKGLTSLFPVN 289

QY 301 MKKADLPQAHSHIKECKGSPVLSLIGALLDFPNRWEYLLKOLQKQPKRIRKSSSYD 360
 DB 290 MKKADLPQAHSHIKECKGSPVLSLIGALLDFPNRWEYLLKOLQKQPKRIRKSSSYD 349
 QY 361 YEALDEAMSISVEMLRDIDKYITDLSILQKQKVPKTVLCILWDMETEVEDILQEFVN 420
 DB 350 YEALDEAMSISVEMLRDIDKYITDLSILQKQKVPKTVLCILWDMETEVEDILQEFVN 409
 QY 421 KSLFLCDRNGKSPRYLLHDLQVDFLTKNCSQLODLHKKIITQFORYHQPTLSPPDEDC 480
 DB 410 KSLFLCDRNGKSPRYLLHDLQVDFLTKNCSQLODLHKKIITQFORYHQPTLSPPDEDC 469
 QY 481 MYWYNFLAYHNASAKMKEICALMPSLDWKAKTELKVGPAHLIHEFVEYHILDEKCAV 540
 DB 470 MYWYNFLAYHNASAKMKEICALMPSLDWKAKTELKVGPAHLIHEFVEYHILDEKCAV 529
 QY 541 SENFOEFLSLNGHLLGQPPFNIVQLGCEPETSVEYQOAKLQAKQVDMGMLYLEWINK 600
 DB 530 SENFOEFLSLNGHLLGQPPFNIVQLGCEPETSVEYQOAKLQAKQVDMGMLYLEWINK 589
 QY 601 KNITNLSRLVVRPHTDAVYHACFSESGQRIASCGADKTLQVFKAEKGKLEIKAHEDV 660
 DB 590 KNITNLSRLVVRPHTDAVYHACFSESGQRIASCGADKTLQVFKAEKGKLEIKAHEDV 649
 QY 661 LCCAFSTDDREIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGS 720
 DB 650 LCCAFSTDDREIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGS 709
 QY 721 SDCFLKLDLNQKCRNMTGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
 DB 710 SDCFLKLDLNQKCRNMTGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 769
 QY 781 INVKOFFNLNDEPQEDMEVIVKCCNSADGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 840
 DB 770 INVKOFFNLNDEPQEDMEVIVKCCNSADGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 829
 QY 841 STIQYCDSPONHLAVNALSOYCVELWNTDSKVDACRGLHLSVHGVMSFPGSSFLTS 900
 DB 830 STIQYCDSPONHLAVNALSOYCVELWNTDSKVDACRGLHLSVHGVMSFPGSSFLTS 889
 QY 901 SDOQIRLWETKVKCKNSAVMLKQEVDFVFOENWVLAVDHIRELOLINGRTQDYLIT 960
 DB 890 SDOQIRLWETKVKCKNSAVMLKQEVDFVFOENWVLAVDHIRELOLINGRTQDYLIT 949
 QY 961 EAQVSCCLSPHLQYIAPGDENGAIELELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 1020
 DB 950 EAQVSCCLSPHLQYIAPGDENGAIELELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 1009
 QY 1021 SSDDAEIQVNWQDLKICIFLRGHQETVKDFRLKNSRLLSWSFDGTVKVWNIIITGNKEKD 1080
 DB 1010 SSDDAEIQVNWQDLKICIFLRGHQETVKDFRLKNSRLLSWSFDGTVKVWNIIITGNKEKD 1069
 QY 1081 FVCHQGTVLSCDISHDATKFSSTADKTKIWSFDLLPLHELGRHNGCVRCSAFVSVDST 1140
 DB 1070 FVCHQGTVLSCDISHDATKFSSTADKTKIWSFDLLPLHELGRHNGCVRCSAFVSVDST 1129
 QY 1141 LLATGDDNGEIRIWNVNGELHLICAPLSEGAATHGWWTDLCFSPDGKMLISAGGYIK 1200
 DB 1130 LLATGDDNGEIRIWNVNGELHLICAPLSEGAATHGWWTDLCFSPDGKMLISAGGYIK 1189
 QY 1201 WNVVGTGSSQTFYNGNKLKIHVSDFKTVYVNDNLGILYLQLE 1248
 DB 1190 WNVVGTGSSQTFYNGNKLKIHVSDFKTVYVNDNLGILYLQLE 1237

RESULT 14

AAW91072

ID AAW91072 standard; protein; 1205 AA.

XX AC AAW91072;

XX DT 07-JUN-1999 (first entry)

XX

DE Apoptosis inducer splice variant Apaf-1L.
 XX Apaf-1L; Apaf-1; splice variant; human; caspase-3; activator; Ced-4;
 KW human; apoptosis; programmed cell death; cancer; AIDS;
 KW multiple sclerosis; inflammation; therapy.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 99..109
 FT Region /note= "additional 11 amino acids compared with Apaf-1"
 FT 613..643
 FT Region /note= "putative WD repeat"
 FT 655..685
 FT Region /note= "putative WD repeat"
 FT 692..729
 FT Region /note= "putative WD repeat"
 FT 741..771
 FT Region /note= "putative WD repeat"
 FT 788..825
 FT Region /note= "putative WD repeat"
 FT 928..946
 FT Region /note= "putative WD repeat"
 FT 958..988
 FT Region /note= "putative WD repeat"
 FT 999..1028
 FT Region /note= "putative WD repeat"
 FT 1040..1070
 FT Region /note= "putative WD repeat"
 FT 1082..1112
 FT Region /note= "putative WD repeat"
 FT 1119..1160
 FT Region /note= "putative WD repeat"
 FT
 XX WO9855615-A1.
 XX
 XX 10-DEC-1998.
 XX
 XX 05-JUN-1998; 98WO-US011773.
 XX
 XX 05-JUN-1997; 97US-0048807P.
 XX 07-AUG-1997; 97US-0055258P.
 XX
 XX (GETH) GENENTECH INC.
 XX (TEXA) UNIV TEXAS.
 XX
 XX Zou H, Henzel WJ, Wang X;
 XX
 XX WPI; 1999-080828/07.
 XX N-PSDB; AAV84798.
 XX
 XX New Apaf-1 protein for regulating apoptosis - is a human homologue of Ced
 XX -4 and an activator of caspase-3; for screening agents that regulate
 XX apoptosis and for treating cancer, AIDS.
 XX
 XX Claim 4; Fig 17; 134pp; English.
 XX
 XX This polypeptide comprises Apaf-1L, a splice variant of Apaf-1 (see
 XX AAW91071), a novel human protein that is believed to be a mammalian
 XX homologue of the Caenorhabditis elegans Ced-4 apoptosis protein. Apaf-1L
 XX cDNA (see AAV84798) was produced from HeLa cell mRNA. The encoded
 XX polypeptide contains 11 additional amino acids compared to Apaf-1, but
 XX retains the structural and functional features of Apaf-1. It is a 130 kDa
 XX polypeptide containing a Ced-3 homologous N-terminal domain, followed by
 XX a Ced-4 homologous domain and multiple C-terminal WD-40 repeats. It forms
 XX a complex with cytochrome-3 and activates caspase-3 in the apoptotic
 XX pathway. Apaf-1L can be produced by culturing cells transformed or
 XX transfected with a vector containing Apaf-1L nucleic acid. Apaf-1
 XX polypeptides, including Apaf-1L, can be used to induce apoptosis, and the
 XX polypeptides, genes and antibodies are especially useful for screening
 XX and identifying cells signalled for apoptosis. Agonistic antibodies can
 XX be used to stimulate apoptosis in cancer cells, and antagonistic
 XX antibodies can be used to block excessive apoptosis or to block the

CC autoimmune and antiinflammatory activities of Apaf-1, in conditions such
 CC as AIDS and multiple sclerosis. Apaf-1 polypeptides and nucleic acids can
 CC be used to screen for agents which modulate apoptosis. Knockout animals
 CC are also provided
 XX

SQ Sequence 1205 AA;

Query Match 96.0%; Score 6352.5; DB 2; Length 1205;

Best Local Similarity 96.6%; Pred. No. 0;

Matches 1205; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

QY	1	MDAKARNCLLOHREALEKDITKSYIMDHMISDGLTISEEEKVRNEPTQOQRAAMLIKMI	60
DB	1	MDAKARNCLLOHREALEKDITKSYIMDHMISDGLTISEEEKVRNEPTQOQRAAMLIKMI	60
QY	61	LKKONDSYVSYFNALLHGYKDIALLLHDGIPVSSSGKDSVSGITSVYRVTLCEGGVP	120
DB	61	LKKONDSYVSYFNALLHGYKDIALLLHDGIPVSSSGKDSVSGITSVYRVTLCEGGVP	120
QY	121	QRPVVFVTRKLVNAIQKLSKLGEPGWTHMGAGCKSVLAABAVRDSHLSLEGCPFG	180
DB	121	QRPVVFVTRKLVNAIQKLSKLGEPGWTHMGAGCKSVLAABAVRDSHLSLEGCPFG	180
QY	181	GVHWVSVCKDQSGLLMKLQNLCTELDQDESQRLPLNIEAKDRLILMLRKHPRLSLL	240
DB	181	GVHWVSVCKDQSGLLMKLQNLCTELDQDESQRLPLNIEAKDRLILMLRKHPRLSLL	240
QY	241	ILDDVWDSWLVKAFDSQOQIILLTRDKSVDSVMGPKVVPVSSSLGKEGLEILSLFVN	300
DB	241	ILDDVWDSWLVKAFDSQOQIILLTRDKSVDSVMGPKVVPVSSSLGKEGLEILSLFVN	300
QY	301	MKKADLPQAHSIIEKCGSPVWLSLIGALLRDPFNRMWYLLKQLQNKQKRIKSSSYD	360
DB	301	MKKADLPQAHSIIEKCGSPVWLSLIGALLRDPFNRMWYLLKQLQNKQKRIKSSSYD	360
QY	361	YEALDEAMSIIVEMLRDIDKYITDLSILOKDKVPTKVLCLMDMETEEVEDILOQFVN	420
DB	361	YEALDEAMSIIVEMLRDIDKYITDLSILOKDKVPTKVLCLMDMETEEVEDILOQFVN	420
QY	421	KSLFCDNRGKSFYYLHDQVDFTEKNCQLODLHKKIITOFORYHOPHTLSPDQDC	480
DB	421	KSLFCDNRGKSFYYLHDQVDFTEKNCQLODLHKKIITOFORYHOPHTLSPDQDC	480
QY	481	MYTNFLAYHMASAKWHKELCALMFSLDWIKAKTELVGPAHLIHEFVYRHLIDKCAV	540
DB	481	MYTNFLAYHMASAKWHKELCALMFSLDWIKAKTELVGPAHLIHEFVYRHLIDKCAV	540
QY	541	SENFOEFLSLNGHLLGROPFPNIVOLGLCEPETSIVYQOAKLQAKQEVNDGMLYLEWINK	600
DB	541	SENFOEFLSLNGHLLGROPFPNIVOLGLCEPETSIVYQOAKLQAKQEVNDGMLYLEWINK	600
QY	601	KNITNLSRLVVRPHTDVYHACFSDGQRIASCADKTLQVFKETGKLEIKAHEDV	660
DB	601	KNITNLSRLVVRPHTDVYHACFSDGQRIASCADKTLQVFKETGKLEIKAHEDV	660
QY	661	LCCAFSTDDRFTATCSVDKVKVIMNSMTGELVHTYDEHSEOVNCCHFTNSHHLLATGS	720
DB	661	LCCAFSTDDRFTATCSVDKVKVIMNSMTGELVHTYDEHSEOVNCCHFTNSHHLLATGS	720
QY	721	SDCFLLKLDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS	780
DB	721	SDCFLLKLDLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS	780
QY	781	INVKOFFLNLDPQEDMEVIVKCCSWSDAGRIWAAKNKILFLDIHTSGLLGEIHTGHH	840
DB	781	INVKOFFLNLDPQEDMEVIVKCCSWSDAGRIWAAKNKIF-----822	822
QY	841	STIQCDPSQNHVALVALLSQVCELTWNTDGRSKVADCRGHLWSVHGMFSPDGSFSLTS	900
DB	823	-----LWNTDGRSKVADCRGHLWSVHGMFSPDGSFSLTS	857
QY	901	SDDOTIRLWETKVKCKNSAVMLKQEVVVFQENVMVLAVDHRLQLINGRTGQIDYLT	960

DB	858	SDDOTIRLWETKVKCKNSAVMLKQEVVVFQENVMVLAVDHRLQLINGRTGQIDYLT	917
QY	961	EAQVSCCLSPHLQYIAFGDENGAIIEILELVNNRIFQSRFQHKKTVMHIQFTADEKTLIS	1020
DB	918	EAQVSCCLSPHLQYIAFGDENGAIIEILELVNNRIFQSRFQHKKTVMHIQFTADEKTLIS	977
QY	1021	SSDDAEIOVWVWQDKCIFLRGHQETVKDFRLLKNSRLLSWSFDGTVKVMNIITGNKEKD	1080
DB	978	SSDDAEIOVWVWQDKCIFLRGHQETVKDFRLLKNSRLLSWSFDGTVKVMNIITGNKEKD	1037
QY	1081	FVCHQGTVLSGDISHDATKFSSTADKTAKTWSFDLPLPLHLRGNHGCVRCSAFSVOST	1140
DB	1038	FVCHQGTVLSGDISHDATKFSSTADKTAKTWSFDLPLPLHLRGNHGCVRCSAFSVOST	1097
QY	1141	LLATGDDNGEIRIMNVSNGELHLHLCAPLSEGAATHGCVMTDLCFSPDGKMLISAGGIK	1200
DB	1098	LLATGDDNGEIRIMNVSNGELHLHLCAPLSEGAATHGCVMTDLCFSPDGKMLISAGGIK	1157
QY	1201	WNVVVTGESSQTFYNTGNTLNKKIHVSPDFKTYVTVDNIGILYILQTL	1248
DB	1158	WNVVVTGESSQTFYNTGNTLNKKIHVSPDFKTYVTVDNIGILYILQTL	1205

RESULT 15

AA97637

ID AA97637 standard; protein; 1205 AA.

XX AA97637;

XX 20-APR-2001 (first entry)

XX Apaf-1L protein sequence.

XX Apaf1; CED-4 homologue; cytochrome c-dependent activation; caspase-3;
 KW apoptosis signalling pathway; cancer; autoimmune disease; variant;
 KW hereditary disease; Apaf-1L.

XX Homo sapiens.

XX WO200100827-A1.

XX 04-JAN-2001.

XX 30-JUN-2000; 2000WO-US018039.

XX 30-JUN-1999; 99US-0141718P.

XX (UNMI) UNIV MICHIGAN.

XX Nunez G, Hu Y;

XX WPI; 2001-112454/12.

XX N-FSDB; AAA91116.

PT Compositions for screening apoptosis pathway agonists and antagonists,
 PT useful for the treatment and diagnosis of cancer and autoimmune diseases,
 PT comprises new splice variants of wild-type Apaf-1 gene.

PS Disclosure; Fig 26; 101pp; English.

XX This sequence represents an Apaf-1 variant of the invention. The variants
 CC of the invention are all splice variants of a wild-type Apaf-1 gene (a
 CC mammalian homolog of CED-4 that participates in cytochrome c-dependent
 CC activation of caspase-3). The DNA is useful in screening assays for
 CC identifying apoptosis signalling pathway (antagonists, which are in turn
 CC useful as potential therapeutics and diagnostic or prognostic tools for
 CC diverse types of cancers, autoimmune diseases and hereditary diseases and
 CC other members of the signalling pathway, i.e. their substrates or
 CC ligands. The protein encoded by the Apaf-1 variant genes are useful in a
 CC cell-free assay system and the antibody generated to the translation
 CC product are used in immunoprecipitation assays to isolate new Apaf-1
 CC pathway constituents or their natural mutants

XX	Sequence 1205 AA;									
SQ	Query Match 96.0%; Score 6352.5; DB 4; Length 1205;									
	Best Local Similarity 96.6%; Pred. No. 0;									
	Matches 1205; Conservative 0; Mismatches 0; Indels 43; Gaps 1;									
QY	1	MDAKARNCLLQHRRALEKDIKTSYVNDHMSDGLFTISEEEKVNEPTQOORAMLIKMI	60							
DB	1	MDAKARNCLLQHRRALEKDIKTSYVNDHMSDGLFTISEEEKVNEPTQOORAMLIKMI	60							
QY	61	LKXNDNSVSVFNALLHEGKQIAALLHDIPIVWSSSKDSVSGITSYVTVLCEGGVP	120							
DB	61	LKXNDNSVSVFNALLHEGKQIAALLHDIPIVWSSSKDSVSGITSYVTVLCEGGVP	120							
QY	121	QRPVVFTRKLVNAIOOKLSKLKGBEGWTHGAGCKSVLAFAVRDHSILKCPFG	180							
DB	121	QRPVVFTRKLVNAIOOKLSKLKGBEGWTHGAGCKSVLAFAVRDHSILKCPFG	180							
QY	181	GVHWSVGQKSGLLMKLQNLCTRLDQDESQRPLNIEAKDRILRLMLRKHPRSL	240							
DB	181	GVHWSVGQKSGLLMKLQNLCTRLDQDESQRPLNIEAKDRILRLMLRKHPRSL	240							
QY	241	ILDDVMDSVLKAQDSQCOILLTRDKSVTDSVMGPKYVVPVSSSLGKEKLEILSLFVN	300							
DB	241	ILDDVMDSVLKAQDSQCOILLTRDKSVTDSVMGPKYVVPVSSSLGKEKLEILSLFVN	300							
QY	301	MKKADLPQAHSLIKECKGSPVSVSLIGALLRDPFNWMEYLLKOLNKKQFKRIRKSSSYD	360							
DB	301	MKKADLPQAHSLIKECKGSPVSVSLIGALLRDPFNWMEYLLKOLNKKQFKRIRKSSSYD	360							
QY	361	YEALDEAMSISVEMLRREDIKDYVTDLSILQKQVPTKVLICILMDMETEEVEDILOEFVN	420							
DB	361	YEALDEAMSISVEMLRREDIKDYVTDLSILQKQVPTKVLICILMDMETEEVEDILOEFVN	420							
QY	421	KSLLFCDRNGKSFYLLHDLQVDFLTKNCSQDLHKKIITQFORVHQPHLSPQEDC	480							
DB	421	KSLLFCDRNGKSFYLLHDLQVDFLTKNCSQDLHKKIITQFORVHQPHLSPQEDC	480							
QY	481	MYWYFLAYHMASAKMHELCAALMFSIDWKAKTELVPAPHLIHEFVEYRHILDEKCAV	540							
DB	481	MYWYFLAYHMASAKMHELCAALMFSIDWKAKTELVPAPHLIHEFVEYRHILDEKCAV	540							
QY	541	SENFQEFISLNGHLLGRPPFNIVOLGLCEPETSEVYQAKLOAKQEVNDGMLYLEWINK	600							
DB	541	SENFQEFISLNGHLLGRPPFNIVOLGLCEPETSEVYQAKLOAKQEVNDGMLYLEWINK	600							
QY	601	KNITNLSRLVVRPHTDVYHACFSEDCQRTASCADKTLQVFKATGKLEIKAHEDV	660							
DB	601	KNITNLSRLVVRPHTDVYHACFSEDCQRTASCADKTLQVFKATGKLEIKAHEDV	660							
QY	661	LCCAFSTDDRFATCSVDKKYIWSMTGELVHTYDEHSEQVNCCHFTNSSHHLLATGS	720							
DB	661	LCCAFSTDDRFATCSVDKKYIWSMTGELVHTYDEHSEQVNCCHFTNSSHHLLATGS	720							
QY	721	SDCFLKLDLNOKECRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS	780							
DB	721	SDCFLKLDLNOKECRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS	780							
QY	781	INVQKFFNLNLEDPQEDMEVIKCCSWSDGARIWAAKNKIFLDIHTSGLLGEIHTGHH	840							
DB	781	INVQKFFNLNLEDPQEDMEVIKCCSWSDGARIWAAKNKIFLDIHTSGLLGEIHTGHH	840							
QY	841	STIQYCDPSPQNHVALVSQYCVELWNTDTSRSKVADCRGHLSSVHGVWVSPDGSFLLTS	900							
DB	841	STIQYCDPSPQNHVALVSQYCVELWNTDTSRSKVADCRGHLSSVHGVWVSPDGSFLLTS	900							
QY	901	SDDOTIRLWETKVCNSAVMLKQEVDFVFOENEVWVLAVDHRLQLINGRTGQIDYLT	960							
DB	901	SDDOTIRLWETKVCNSAVMLKQEVDFVFOENEVWVLAVDHRLQLINGRTGQIDYLT	960							
QY	961	EAQVSCCLSPHLQVIAFGDENGAIETLELVNRRIFQSFQHKKTWHIQTADADEKTLIS	1020							
DB	961	EAQVSCCLSPHLQVIAFGDENGAIETLELVNRRIFQSFQHKKTWHIQTADADEKTLIS	1020							

Db	918	EAQVSCCLSPHLQVIAFGDENGAIETLELVNRRIFQSFQHKKTWHIQTADADEKTLIS	977
QY	1021	SSDDAEIQVWVWQDKCIFIPLRGHQTVDKDFRLLKNSRLLSWSFDGTVKVWNIITGNKEKD	1080
Db	978	SSDDAEIQVWVWQDKCIFIPLRGHQTVDKDFRLLKNSRLLSWSFDGTVKVWNIITGNKEKD	1037
QY	1081	FVCHQGTVLSCDISHDQATKFSSTADKTAKIWSFDLLPLHFLRGHNGCVRCSAFSVDST	1140
Db	1038	FVCHQGTVLSCDISHDQATKFSSTADKTAKIWSFDLLPLHFLRGHNGCVRCSAFSVDST	1097
QY	1141	LLATGDDNGEIRIWNVSNGELLHLCAPLSEEGAAHGGWVTDLCFSPDGMKLIISAGGYIK	1200
Db	1098	LLATGDDNGEIRIWNVSNGELLHLCAPLSEEGAAHGGWVTDLCFSPDGMKLIISAGGYIK	1157
QY	1201	WNVVVTGESSQTFYVNGTNLKKIHVSPOFKTVTVVDNLGILYILOTLLE	1248
Db	1158	WNVVVTGESSQTFYVNGTNLKKIHVSPOFKTVTVVDNLGILYILOTLLE	1205

Search completed: January 27, 2005, 18:36:24
Job time : 184 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 18:30:10 ; Search time 52 Seconds
(without alignments)
2309.202 Million cell updates/sec

Title: US-10-646-396-2
Perfect score: 6619
Sequence: 1 MDKARNCLLQHREALSKDI.....FKTVVVDNLGILYILQTL 1248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6282	94.9	1194	2 T03818	apoptotic proteina
2	732	11.1	1227	2 A21810	WD-40 repeat prote
3	728.5	11.0	1258	2 A12155	WD-repeat protei
4	691	10.4	1526	2 AC2239	WD-40 repeat prote
5	677.5	10.2	1711	2 AD1842	WD-40 repeat prote
6	672	10.2	1683	2 AF2071	WD-40 repeat prote
7	643.5	9.7	1356	2 T18521	beta transducin-li
8	626	9.5	1693	2 S76086	beta transducin-li
9	614.5	9.3	934	2 AG1889	WD-40 repeat prote
10	611.5	9.2	1551	2 AB2410	WD-repeat protei
11	610	9.2	1708	2 AE1866	WD-40 repeat prote
12	600.5	9.1	1189	2 A12493	WD-repeat protei
13	580.5	8.8	1189	2 AH2154	WD-repeat protei
14	561.5	8.5	1747	2 AC1842	WD-40 repeat prote
15	464.5	7.0	1329	2 AB1901	WD-repeat containi
16	441.5	6.7	1101	2 T26919	hypothetical prote
17	423	6.4	2629	2 T32735	telomerase-associ
18	399.5	6.0	2629	2 T30987	telomerase-associ
19	396	6.0	1049	2 T42045	beta transducin-li
20	394	6.0	777	2 T41075	hypothetical WD-re
21	391.5	5.9	1191	2 S76414	beta transducin-li
22	385	5.8	677	2 AE1861	serine/threonine k
23	370	5.6	317	2 T46032	WD-40 repeat regul
24	369	5.6	786	2 AG2375	WD-40 repeat prote
25	365	5.5	676	2 AH2195	hypothetical prote
26	355	5.4	876	2 T51507	WD40-repeat protei
27	338.5	5.1	589	2 AG2400	WD-repeat protei
28	326	4.9	559	2 AB2202	hypothetical prote
29	322.5	4.9	515	2 S19487	hypothetical prote

RESULT 1

T03818

apoptotic proteinase-activating factor 1 - human

N:Alternate names: APAF-1 protein; protein KIAA0413

C:Species: Homo sapiens (man)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T03818; T00053

R;Zou, H.; Henzel, W.J.; Liu, X.S.; Lutschg, A.; Wang, X.D.

Cell 90, 405-413, 1997

A:Title: Apaf-1, a human protein homologous to C. elegans CED-4, participates in cytochr

A:Reference number: Z15099; MUID:97410306; PMID:9267021

A;Accession: T03818

A;Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A;Residues: 1-1194 <ZOU>

A;Cross-references: UNIPROT:O14727; EMBL:AF013263; NID:g2330014; PIDN:AAC51678.1; PID:g2

A;Experimental source: cell line HeLa S3

R;Ishikawa, K.; Nagase, T.; Nakajima, D.; Seki, N.; Ohira, M.; Miyajima, N.; Tanaka, A.,

submitted to the EMBL Data Library, October 1997

A;Description: Prediction of the coding sequences of unidentified human genes. VIII. The

A;Reference number: Z14080

A;Accession: T00053

A;Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A;Residues: 374-812, 'FDIHTSGLLGHEIHTGHSTIQYCDSPQNHVAVVALSYQYCVEL', 813-1186, 'I', 1188-1194

A;Cross-references: EMBL:AB007873; NID:dl179749; PIDN:BAA24843.1; PID:dl025765

A;Experimental source: brain

C;Genetics:

A;Gene: Apaf-1

A;Note: KIAA0413

C;Function:

A;Description: binds and hydrolyzes ATP or dATP to ADP or dADP, respectively; participat

Query Match 94.9%; Score 6282; DB 2; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches - 0; Indels 54; Gaps 2;

Qy	1	MDAKARNCLLQHREALKDIKTSYIMDHMSDGLTITSEEEKVNRNEPTQQQRAAMLIKMI	60
Db	1	MDAKARNCLLQHREALKDIKTSYIMDHMSDGLTITSEEEKVNRNEPTQQQRAAMLIKMI	60
Qy	61	LKKNDNSVSYFNALLHEGYKDLAALLHDGIPVVSSSSKDSVSGITTSYVTVLCEGGVP	120
Db	61	LKKNDNSVSYFNALLHEGYKDLAALLHDGIPVVSSSS-----VTVLCEGGVP	109
Qy	121	QRPVVFVTRKLVNAIOQKSLKGEPEGWVTHGVAGCGKSVLAEEAVDRDHSLLGCGPPG	180
Db	110	QRPVVFVTRKLVNAIOQKSLKGEPEGWVTHGVAGCGKSVLAEEAVDRDHSLLGCGPPG	169
Qy	181	GVHWSVGKQDKSLMLKQLNLCITRLDQDSFSQRLPLNIEAKDRLILMLRKHPRSL	240
Db	170	GVHWSVGKQDKSLMLKQLNLCITRLDQDSFSQRLPLNIEAKDRLILMLRKHPRSL	229

30	320.5	4.8	304	2	AG1837	WD-40 repeat prote
31	318	4.8	817	2	S51445	probable membrane
32	311.5	4.7	410	2	S48052	platelet-activati
33	311.5	4.7	981	2	T18234	beta transducin ho
34	309	4.7	318	2	S11904	GTP-binding regula
35	307	4.6	409	2	S36113	LIS-1 protein - hu
36	306	4.6	507	2	T38653	trp-asp repeat pro
37	304	4.6	922	2	T40372	trp asp repeat pro
38	302	4.6	342	2	AE2490	WD-repeat protei
39	302	4.6	357	2	A12099	WD-40 repeat prote
40	299	4.5	333	2	G85034	probable WD-repeat
41	297.5	4.5	302	2	T41148	trp-asp repeat con
42	297.5	4.5	640	2	S49932	MET30 protein - ye
43	297	4.5	586	2	T38992	WD-40 repeat regul
44	296.5	4.5	473	2	T3805	hypothetical prote
45	294	4.4	316	2	S57839	CPC2 protein - Neu

ALIGNMENTS

QY 241 ILDDVMDSWLKAFDSQCOILLTRDKSVTDSVMGPKYVVPVSSSLGKGLBILSLFVN 300
 DB 230 ILDDVMDSWLKAFDSQCOILLTRDKSVTDSVMGPKYVVPVSSSLGKGLBILSLFVN 289
 QY 301 MKKADLPQAHSIIECKGSPVLSLIGALLRDPFNRWEYLLKQNKQPKRIRKSSSYD 360
 DB 290 MKKADLPQAHSIIECKGSPVLSLIGALLRDPFNRWEYLLKQNKQPKRIRKSSSYD 349
 QY 361 YEALDEAMSISVEMREDIKDYITDLSILQKQVPTKVLGILMDMETEEVEDILOEFVN 420
 DB 350 YEALDEAMSISVEMREDIKDYITDLSILQKQVPTKVLGILMDMETEEVEDILOEFVN 409
 QY 421 KSLIFCDRNGKSFYRLHDLQVDFLEKNCOSQLDHLKKIITQORVHOPHTLSPOEDC 480
 DB 410 KSLIFCDRNGKSFYRLHDLQVDFLEKNCOSQLDHLKKIITQORVHOPHTLSPOEDC 469
 QY 481 MYWYNFLAYHWSAKMHEKCALMFSLDWKATLAVGPANLHFEFVEYRHILDEKCAV 540
 DB 470 MYWYNFLAYHWSAKMHEKCALMFSLDWKATLAVGPANLHFEFVEYRHILDEKCAV 529
 QY 541 SENFOEFLSLNGHILGROPENIVOLGCEPETSEVYOAKLOAKQVDMGMLYLEWINK 600
 DB 530 SENFOEFLSLNGHILGROPENIVOLGCEPETSEVYOAKLOAKQVDMGMLYLEWINK 589
 QY 601 KMITNLRLVPRDTHDVAHYHACSEDOQRIASCGADKTLQVFKATGEKLELKAHEDEV 660
 DB 590 KMITNLRLVPRDTHDVAHYHACSEDOQRIASCGADKTLQVFKATGEKLELKAHEDEV 649
 QY 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
 DB 650 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 709
 QY 721 SDFCLKLDLQKQCRNMTGHTNSVNHCRFPDOKLLASCADGTLKLDATSANERKS 780
 DB 710 SDFCLKLDLQKQCRNMTGHTNSVNHCRFPDOKLLASCADGTLKLDATSANERKS 769
 QY 781 INVKOFFNLNLEDPQEDMEVIVKCCSWADGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 840
 DB 770 INVKOFFNLNLEDPQEDMEVIVKCCSWADGARIWAAKNKIFLFDIHTSGLLGEIHTGHH 811
 QY 841 STIQYCDSPQNLAVVALSYCVELWNTDSRSKADCRHLSVHVGWVSPDGSFLTS 900
 DB 812 STIQYCDSPQNLAVVALSYCVELWNTDSRSKADCRHLSVHVGWVSPDGSFLTS 846
 QY 901 SDDQITRLWETKVKCKNSAVMLKQEDVVDVVOENWVWLAVDHRLQOLINGRTQIDYLT 960
 DB 847 SDDQITRLWETKVKCKNSAVMLKQEDVVDVVOENWVWLAVDHRLQOLINGRTQIDYLT 906
 QY 961 EAQVSCCLSPHLQYIAFGDENGAIETELVNNRIFQSEFQHKTVWHIQTADKTLIS 1020
 DB 907 EAQVSCCLSPHLQYIAFGDENGAIETELVNNRIFQSEFQHKTVWHIQTADKTLIS 966
 QY 1021 SDDAEIQVNNQDLKCIPLAGHQTQVDFRLLKNSRLLSFGDTGVKWNITGNKEVD 1080
 DB 967 SDDAEIQVNNQDLKCIPLAGHQTQVDFRLLKNSRLLSFGDTGVKWNITGNKEVD 1026
 QY 1081 FVCHQGTVLSCDISHDATKFSSTSAKTAKIWSFDLLPLHELGHNGCVRCSAFSDST 1140
 DB 1027 FVCHQGTVLSCDISHDATKFSSTSAKTAKIWSFDLLPLHELGHNGCVRCSAFSDST 1086
 QY 1141 LLATGDDNGEIRIWNVNSGELLHLCAPLSEGAATHGGVWTDLCFSPDGKMLISAGGIYK 1200
 DB 1087 LLATGDDNGEIRIWNVNSGELLHLCAPLSEGAATHGGVWTDLCFSPDGKMLISAGGIYK 1146
 QY 1201 WNNVVTGSSQFTYNTGNLKKIHVSDFKTYVVDNLGILYIQTLE 1248
 DB 1147 WNNVVTGSSQFTYNTGNLKKIHVSDFKTYVVDNLGILYIQTLE 1194

RESULT 2
 AE1810
 WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C;Accession: AE1810
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AE1810
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1227 <KUR>
 A;Cross-references: UNIPROT:Q8Z0R1; GB:BA000019; PIDN:BA077553.1; PID:gl7135007; GSPDB:G
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: alr0029

Query Match 11.1%; Score 732; DB 2; Length 1227;
 Best Local Similarity 22.2%; Pred. No. 1.3e-35;
 Matches 295; Conservative 209; Mismatches 438; Indels 384; Gaps 50;

QY 77 HEGYKDLAALLHDGIPWSSSSGKDSVSGITSYVTVLCEGVPQRPVVPVTRKKLVAI 136
 DB 45 HSG-AETAAKLDISQAVRKLGE-----SY-RKLGIEG-----KGNKINGL 85
 QY 137 QOKLSK-----LKGEPGW-----VTIHG 154
 DB 86 ROKLYEQVLSYQHPVSSSEDGAEVDVAGFRGRKEPPELLEQWIEGNSNRCRLVAVLG 145
 QY 155 MAGCKSVLAEEAVR-----DHSLLGCGFGGVWVSVGKQKSG-LLMKLONLCTRLD 207
 DB 146 MGGTGKTVLAAMGAKKQVKEFDYLI-----WRLSNAPSLGDLITQLLRLFLANDN 195
 QY 208 QDEFSQRLPINTEEAKDRILRLMKRHPRLSLLIDVW-----DSW-----249
 DB 196 END-----LTDTDNNKIVRLDVLKRCGLIILDNVESVLSRSGEGKQEWAGDYQPGY 248
 QY 250 -----VLKAFDSQCOILLTRDKSVTDSVMGPKYVVPVSSSLGKGLBILSL-FV 299
 DB 249 ENYGFLPKVAEASHESC-LLLTSREK-----PKEVAAL-----GKNLPVKVQLSSL 296
 QY 300 NKKKAD-----LPEQAHSIIECKGSPVLSLIGALLRD-FPNRWYYLQK-----344
 DB 297 NLAEAREILLDKGYCTDEQDELVRYSNGNPLAKIVATTVVELFSNNISEFLTQHQE 356
 QY 345 -----LQNKQFKR-----IRKSSSYDEALDEAMSISVEMLEKEDIKYITDLSILOXD 392
 DB 357 SAVYGDITLLKQOFQRLSELEKVMYSLGANREYVS-----FRELKDDMLTTES-----406
 QY 393 VKVPTKVLCLMDMETEEVEDILOEFV-----NKSLLPCDRNGKSFYRLHDLQV 442
 DB 407 ---PIKVM-----EALSLRLSLIEKASPTLIEKASSTQGEKEAESKFGLESVVM 455
 QY 443 DFLTEKNCOSQLDHLKKIITQORVHOPHTLSPOEDCQWYNFL-AYHWSAKMHEKLC 501
 DB 456 EYTH-----KFIENSLEEFSTQKUL-----DFINTYPLMKAR-----488
 QY 502 ALMPSLDWIKATVELGPAHLHFEFVEYRHILDEKDCAVSENFOEFLSLNG-----HL---554
 DB 489 ---SLDVIROIQE-----RLILEPVR-----QKLNIFGTLELHLRR 523
 QY 555 ---IGRQPPP-----NIV-----QLGLCE-PETSEV-----YQOAKL 582
 DB 524 MLGTLOKEPLPKGYAGNLINLLRLQLDLKIPDESPIJLSGRDFFSGLTITWQAYFKEVKL 583
 QY 583 Q-----AKQEVNDGMLYLEWINK-----KNITNLRLVVRP 613
 DB 584 KETIFANSDL-TGSVFTEFTNNSVVSVKFSPDGKYFATGLMNGEIRLWQTSNDKQLRIYKG 642
 QY 614 HTDAVTHACFSEDQRIASCGADKTLQVFKATGEKLELKAHEDEVLCAPSTDDRFIA 673
 DB 643 HTAWWAFAPSPDRMLASGASDSTIKLMDVHTGECLEKTLKNTNKNVSVAFSPDGRILA 702

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 18:40:29 ; Search time 172 Seconds
(without alignments)
2621.449 Million cell updates/sec

Title: US-10-646-396-2

Perfect score: 6619

Sequence: 1 MDAKARNCLLOHREALEKDI.....PKTVTVVDNLGILVILQTL 1248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6619	100.0	1248	17	US-10-482-952-16
2	6352.5	96.0	1205	9	US-09-876-667-16
3	6282	94.9	1194	9	US-09-876-667-2
4	6282	94.9	1194	14	US-10-141-618-10
5	6003.5	90.7	1199	15	US-10-112-944-273
6	5933.5	89.6	1526	15	US-10-112-944-272
7	1589	24.0	308	13	US-10-014-269-31
8	1589	24.0	308	13	US-10-002-974-31
9	1589	24.0	308	14	US-10-314-506-31
10	748	11.3	142	14	US-10-001-254-29
11	711	10.7	529	14	US-10-369-493-12913
12	690.5	10.4	610	14	US-10-369-493-20185
13	688	10.4	1005	14	US-10-369-493-20242

14	674	10.2	1140	14	US-10-369-493-19053	Sequence 19053, A
15	672	10.2	1136	14	US-10-369-493-19046	Sequence 19046, A
16	652	9.9	579	14	US-10-369-493-18893	Sequence 18893, A
17	643.5	9.7	1356	13	US-10-077-111-10	Sequence 10, Appl
18	632.5	9.6	1118	14	US-10-369-493-18845	Sequence 18845, A
19	606.5	9.2	1155	14	US-10-369-493-19869	Sequence 19869, A
20	604.5	9.1	1117	14	US-10-369-493-18965	Sequence 18965, A
21	577	8.7	608	14	US-10-369-493-20224	Sequence 20224, A
22	567	8.6	613	14	US-10-369-493-11181	Sequence 11181, A
23	555.5	8.4	606	14	US-10-369-493-18907	Sequence 18907, A
24	552	8.3	625	14	US-10-369-493-20066	Sequence 20066, A
25	526.5	8.0	656	14	US-10-369-493-20247	Sequence 20247, A
26	526.5	8.0	679	14	US-10-369-493-20059	Sequence 20059, A
27	522.5	7.9	478	14	US-10-369-493-18905	Sequence 18905, A
28	496.5	7.5	574	14	US-10-369-493-20058	Sequence 20058, A
29	483	7.3	584	14	US-10-369-493-19944	Sequence 19944, A
30	476	7.2	93	13	US-10-014-269-28	Sequence 28, Appl
31	476	7.2	93	13	US-10-002-974-28	Sequence 28, Appl
32	476	7.2	93	14	US-10-314-506-28	Sequence 28, Appl
33	461	7.0	90	14	US-10-156-733-4	Sequence 4, Appl
34	455	6.9	89	9	US-09-841-739-16	Sequence 16, Appl
35	455	6.9	89	9	US-09-931-071-8	Sequence 8, Appl
36	444.5	6.7	429	14	US-10-449-315-16	Sequence 16, Appl
37	444.5	6.7	429	14	US-10-369-493-11348	Sequence 11348, A
38	430	6.5	1102	14	US-10-369-493-19941	Sequence 19941, A
39	423	6.4	2629	14	US-10-295-681-54	Sequence 54, Appl
40	423	6.4	2629	14	US-10-295-681-55	Sequence 55, Appl
41	423	6.4	2629	14	US-10-295-681-61	Sequence 61, Appl
42	410	6.2	2630	15	US-10-334-143-41	Sequence 41, Appl
43	368.5	5.6	559	14	US-10-369-493-20028	Sequence 20028, A
44	366	5.5	891	16	US-10-437-363-153705	Sequence 153705, A
45	358	5.4	280	14	US-10-369-493-19899	Sequence 19899, A

ALIGNMENTS

RESULT 1

US-10-482-952-16
; Sequence 16, Application US/10482952
; Publication No. US20040254136A1
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Columbia University in the City of New York, et al.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES AND RELATED METHODS FOR REGULATING CBI
; FILE REFERENCE: 0575/64078-A-PCT/JPW/ANX
; CURRENT APPLICATION NUMBER: US/10/482,952
; CURRENT FILING DATE: 2004-01-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 16
; LENGTH: 1248
; TYPE: PRT
; ORGANISM: human
US-10-482-952-16

Query Match 100.0%; Score 6619; DB 17; Length 1248;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MDAKARNCLLOHREALEKDIKTSYIMDHIMISDGFITISEEEKVNRNEPTQQRAAMLIKMI	60
Db	1	MDAKARNCLLOHREALEKDIKTSYIMDHIMISDGFITISEEEKVNRNEPTQQRAAMLIKMI	60
Qy	61	LKNDNDVSVPYNALLHEGYKDALLHDGIPVYSSSSGKDSVSGITSYVTVLCEGVP	120
Db	61	LKNDNDVSVPYNALLHEGYKDALLHDGIPVYSSSSGKDSVSGITSYVTVLCEGVP	120
Qy	121	QRPVVFVTRKLVNAIQKLSKLGEPGWVTHGMAGCKSVLAEEA VRDHSLLGCFPG	180
Db	121	QRPVVFVTRKLVNAIQKLSKLGEPGWVTHGMAGCKSVLAEEA VRDHSLLGCFPG	180
Qy	181	GVHWSVGCKDQSGLLMKIQNLCTRLDODESFSORPLNIEAKDRLMLRKHPRSLL	240

```

Db 181 GVHWVSVGQDKSGLLMKLQNLCTRLDQDESPQRPLNIEBAKDRILMLRKHPSLL 240
Qy 241 ILDDVMSWLVKAFQSQCIILLTTRDKSVTDSVMGPKYVVPVSSGLGKGLKLEILSLFN 300
Db 241 ILDDVMSWLVKAFQSQCIILLTTRDKSVTDSVMGPKYVVPVSSGLGKGLKLEILSLFN 300
Qy 301 MKKADLPEQAHSIKECKGSLVSLIGALLRDPFNNRWEYILKQKQKPKRIRKSSSYD 360
Db 301 MKKADLPEQAHSIKECKGSLVSLIGALLRDPFNNRWEYILKQKQKPKRIRKSSSYD 360
Qy 361 YEALDEAMSISEMLREDIKDYITDLSILOKOVKVPKVLCTILMDMETEEVEDILOEFVN 420
Db 361 YEALDEAMSISEMLREDIKDYITDLSILOKOVKVPKVLCTILMDMETEEVEDILOEFVN 420
Qy 421 KSLIFCDNRNGKFRYYLHDLQVDFLTKNGSQDLHKKIITQFORVHOPHTLSPQDEDC 480
Db 421 KSLIFCDNRNGKFRYYLHDLQVDFLTKNGSQDLHKKIITQFORVHOPHTLSPQDEDC 480
Qy 481 MYWYNFLAYHMASAKWHELCALMFSLDWIKATELVGPAHLIHEFVEYRHILDEKCAV 540
Db 481 MYWYNFLAYHMASAKWHELCALMFSLDWIKATELVGPAHLIHEFVEYRHILDEKCAV 540
Qy 541 SENFOEFLSLNGHLGROPENIVOLGLCEPETSEVYQOAKLOAKQEVNDGMLYLEWINK 600
Db 541 SENFOEFLSLNGHLGROPENIVOLGLCEPETSEVYQOAKLOAKQEVNDGMLYLEWINK 600
Qy 601 KNTNLRLVVRPHDTAVTHACSEDOQRTASCADKTLQVFKARTGEKLEIKAHEDV 660
Db 601 KNTNLRLVVRPHDTAVTHACSEDOQRTASCADKTLQVFKARTGEKLEIKAHEDV 660
Qy 661 LCCAFSTDDRFTATCSVDKVKIWSMTGELVHTYDEHSEQVNCCHFTNSHLLLATGS 720
Db 661 LCCAFSTDDRFTATCSVDKVKIWSMTGELVHTYDEHSEQVNCCHFTNSHLLLATGS 720
Qy 721 SDCLFLKLDLNOKECRNTMFHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
Db 721 SDCLFLKLDLNOKECRNTMFHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
Qy 781 INVQKPFNLDPQEDMEVIVKCCSWSADGARIMVAANKIPLFDIHTSGLLGEIHTGHH 840
Db 781 INVQKPFNLDPQEDMEVIVKCCSWSADGARIMVAANKIPLFDIHTSGLLGEIHTGHH 840
Qy 841 STTIQYCDSPQNLAVVALSQCVELWNTDSRSKVADCRGHLVWVGVMFSPDGSFLTS 900
Db 841 STTIQYCDSPQNLAVVALSQCVELWNTDSRSKVADCRGHLVWVGVMFSPDGSFLTS 900
Qy 901 SDDQITRLWETKVKCKSAVMKQEVVDVVFQENVMVLAVDHIRRLQILINGRGQIDYLT 960
Db 901 SDDQITRLWETKVKCKSAVMKQEVVDVVFQENVMVLAVDHIRRLQILINGRGQIDYLT 960
Qy 961 EAQVSCCLSLPHLOVIAFGDENGAIETLELVNRRIFQSRFOHKKTVWHIOTFADEKTLIS 1020
Db 961 EAQVSCCLSLPHLOVIAFGDENGAIETLELVNRRIFQSRFOHKKTVWHIOTFADEKTLIS 1020
Qy 1021 SSDDAEIOVWNNWOLDKCIPLRGHOETVKOPRLLKNSRLLSWSDGTWVKNVNIITGNKED 1080
Db 1021 SSDDAEIOVWNNWOLDKCIPLRGHOETVKOPRLLKNSRLLSWSDGTWVKNVNIITGNKED 1080
Qy 1081 FVCHQGTVLSCDISHDATKSSISADTKATKISFDLILPLHELGHNGCVRCGAFSDST 1140
Db 1081 FVCHQGTVLSCDISHDATKSSISADTKATKISFDLILPLHELGHNGCVRCGAFSDST 1140
Qy 1141 LIATGDNGEIRINWNSNGELLHLHCAPLSEGAATHGQWVTDLCFSPDGKMLISAGGIK 1200
Db 1141 LIATGDNGEIRINWNSNGELLHLHCAPLSEGAATHGQWVTDLCFSPDGKMLISAGGIK 1200
Qy 1201 WNNVVTGSSQFTYNGTGNLKKIHVSPDFKTYVTVDNILGILYILOTLE 1248
Db 1201 WNNVVTGSSQFTYNGTGNLKKIHVSPDFKTYVTVDNILGILYILOTLE 1248

```

RESULT 2
US-09-876-667-16

```

; Sequence 16, Application US/09876667
; Patent No. US20020107370A1
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. US20020107370A1west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/876,667
; FILING DATE: 07-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/435,115
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; -TYPE: amino acid
; LENGTH: 1205 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
;
; US-09-876-667-16
;
; Query Match 96.0%; Score 6352.5; DB 9; Length 1205;
; Best Local Similarity 96.6%; Pred. No. 0;
; Matches 1205; Conservative 0; Mismatches 0; Indels 43; Gaps 1;
;
; Qy 1 MDKARNCLLQHREALKDKITSYIMDHMISDGLFTISEBEKVRNEPTQOORAAMLIKMI 60
; Db 1 MDKARNCLLQHREALKDKITSYIMDHMISDGLFTISEBEKVRNEPTQOORAAMLIKMI 60
;
; Qy 61 LKKONDSVSVFYNALHREGYKDLAALLHDGIPVSVSSSGKDSVSGITSYVRTVLCGGVP 120
; Db 61 LKKONDSVSVFYNALHREGYKDLAALLHDGIPVSVSSSGKDSVSGITSYVRTVLCGGVP 120
;
; Qy 121 QRPVVFVTRKKLVNATQOKLSLKGRCGWTHHMGAGCGSVLAARVDRHSLIEGCPG 180
; Db 121 QRPVVFVTRKKLVNATQOKLSLKGRCGWTHHMGAGCGSVLAARVDRHSLIEGCPG 180
;
; Qy 181 GVHWVSVGKQDKSGLLMKLQNLCTRLDQDESPQRPLNIEBAKDRILMLRKHPSLL 240
; Db 181 GVHWVSVGKQDKSGLLMKLQNLCTRLDQDESPQRPLNIEBAKDRILMLRKHPSLL 240
;
; Qy 241 ILDDVMSWLVKAFDQSQCIILLTTRDKSVTDSVMGPKYVVPVSSGLGKGLKLEILSLFN 300
; Db 241 ILDDVMSWLVKAFDQSQCIILLTTRDKSVTDSVMGPKYVVPVSSGLGKGLKLEILSLFN 300
;
; Qy 301 MKKADLPEQAHSIKECKGSLVSLIGALLRDPFNNRWEYILKQKQKPKRIRKSSSYD 360
; Db 301 MKKADLPEQAHSIKECKGSLVSLIGALLRDPFNNRWEYILKQKQKPKRIRKSSSYD 360

```

```

Qy 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKDVKVPVKVLCILMDMETEEVEDILOQFVN 420
Db 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKDVKVPVKVLCILMDMETEEVEDILOQFVN 420
Qy 421 KSLIFCDNRNGSKFRYLLHDQVDFLTKNCSQLODLHKKIITQORVHOPHTLSPDOEDC 480
Db 421 KSLIFCDNRNGSKFRYLLHDQVDFLTKNCSQLODLHKKIITQORVHOPHTLSPDOEDC 480
Qy 481 MYWNLFLAYMASAKHKLALCALMESLDWIKAETELVGPAPLHIEFVYRHLDEKOCV 540
Db 481 MYWNLFLAYMASAKHKLALCALMESLDWIKAETELVGPAPLHIEFVYRHLDEKOCV 540
Qy 541 SENFOEFLSLNGHLGRPFNNIVOLGLCEPETSEVYOAKLOAKQEVNDGMLYLEWINK 600
Db 541 SENFOEFLSLNGHLGRPFNNIVOLGLCEPETSEVYOAKLOAKQEVNDGMLYLEWINK 600
Qy 601 KNIITNLSRLVVRPHTDAVTHACFSEDGQRIASCAGADKTLOVFKAGTGEKLEIKAHDEV 660
Db 601 KNIITNLSRLVVRPHTDAVTHACFSEDGQRIASCAGADKTLOVFKAGTGEKLEIKAHDEV 660
Qy 661 LCCAFSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS 720
Db 661 LCCAFSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS 720
Qy 721 SDCFLKLDLWLNQECRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLMDATSANERKS 780
Db 721 SDCFLKLDLWLNQECRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLMDATSANERKS 780
Qy 781 INVQOFFLNLDEPOEDMEVIVKCCSWADGARIWVAANKIIFLPHIHTSGLLGEIHTGHH 840
Db 781 INVQOFFLNLDEPOEDMEVIVKCCSWADGARIWVAANKIIFLPHIHTSGLLGEIHTGHH 840
Qy 841 STIOYCDPSPQNHAAVALSQCVELWNTDGRSKVADCRGHLVWVHGMVSPDGSSFLT 900
Db 841 STIOYCDPSPQNHAAVALSQCVELWNTDGRSKVADCRGHLVWVHGMVSPDGSSFLT 900
Qy 882 INVQOFFLNLDEPOEDMEVIVKCCSWADGARIWVAANKIIFLPHIHTSGLLGEIHTGHH 882
Db 882 INVQOFFLNLDEPOEDMEVIVKCCSWADGARIWVAANKIIFLPHIHTSGLLGEIHTGHH 882
Qy 901 SDDQIRLWETKVKCKNSAVMLKQEVVVOENEVWVLAVDHIRRLQLINGRTQOIDL 960
Db 901 SDDQIRLWETKVKCKNSAVMLKQEVVVOENEVWVLAVDHIRRLQLINGRTQOIDL 960
Qy 961 EAQVSCCLSLHQLYIAFGDENGAIIELELVNRIQSRFOHKTVMWHIQTADKTLIS 1020
Db 961 EAQVSCCLSLHQLYIAFGDENGAIIELELVNRIQSRFOHKTVMWHIQTADKTLIS 1020
Qy 1021 SDDAEIOVWNNQDCKCIFLGHQETVKDFRLLKNSRLLSVFDGTVKVMNIIITGNKEK 1080
Db 1021 SDDAEIOVWNNQDCKCIFLGHQETVKDFRLLKNSRLLSVFDGTVKVMNIIITGNKEK 1080
Qy 1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLPLPLHELGHNGCVRCSPVSDST 1140
Db 1081 FVCHQGTVLSCDISHDATKFSSTADKTAKIWSFDLPLPLHELGHNGCVRCSPVSDST 1140
Qy 1141 LLATGDDNGEIRIWNVSGELHLHLCAPLSEGAATHGCVWTDLCSPDGKMLISAGGYIK 1200
Db 1141 LLATGDDNGEIRIWNVSGELHLHLCAPLSEGAATHGCVWTDLCSPDGKMLISAGGYIK 1200
Qy 1201 WNVVTGSSOTFYTNNGNLKIHVSDFKTVVTDNLGILYILOTLE 1248
Db 1201 WNVVTGSSOTFYTNNGNLKIHVSDFKTVVTDNLGILYILOTLE 1248
Qy 1248 WNVVTGSSOTFYTNNGNLKIHVSDFKTVVTDNLGILYILOTLE 1248
Db 1248 WNVVTGSSOTFYTNNGNLKIHVSDFKTVVTDNLGILYILOTLE 1248

```

RESULT 3

```

US-09-876-667-2
; Sequence 2, Application US/09876667
; Patent No. US20020107370A1
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAP-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. US20020107370Alwest Center, 90 South Seventh St

```

```

; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/876,667
; FILING DATE: 07-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/435,115
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-876-667-2

```

```

Query Match 94.9%; Score 6282; DB 9; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

Qy 1 MDKARNCCLLOHRALEKDKITSYIMDHMISDGLTITSEEEKVNEPTQOORAAAMLIKMI 60
Db 1 MDKARNCCLLOHRALEKDKITSYIMDHMISDGLTITSEEEKVNEPTQOORAAAMLIKMI 60
Qy 61 LKNDNSVSPYNALLHEGYKDLAALLHDGIPVYSSSSGKDSVSGITSYVTVLCEGVP 120
Db 61 LKNDNSVSPYNALLHEGYKDLAALLHDGIPVYSSSSGKDSVSGITSYVTVLCEGVP 120
Qy 121 QRPVVFVTRKLVNAIOOKLSKLGEPGWVTHGMAGCGKSVLAABAVRDHSLLEGCPFG 180
Db 121 QRPVVFVTRKLVNAIOOKLSKLGEPGWVTHGMAGCGKSVLAABAVRDHSLLEGCPFG 180
Qy 181 GVHWSVKGQDKSGLLMLKQLNCTRLDQDESFSQRLPLNIEAKDRILMLRKHPSLL 240
Db 181 GVHWSVKGQDKSGLLMLKQLNCTRLDQDESFSQRLPLNIEAKDRILMLRKHPSLL 240
Qy 241 ILDDVDSWLVKAFDSQCOILLTTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN 300
Db 241 ILDDVDSWLVKAFDSQCOILLTTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN 300
Qy 301 MKKADLPQAHSIIECKGSPVLSLIGALLRDPFNRWEYLLKQLONKQFKRIRKSSSYD 360
Db 301 MKKADLPQAHSIIECKGSPVLSLIGALLRDPFNRWEYLLKQLONKQFKRIRKSSSYD 360
Qy 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKDVKVPVKVLCILMDMETEEVEDILOQFVN 420
Db 361 YEALDEAMSISVEMLRREDIKDYITDLSILOKDVKVPVKVLCILMDMETEEVEDILOQFVN 420
Qy 421 KSLIFCDNRNGSKFRYLLHDQVDFLTKNCSQLODLHKKIITQORVHOPHTLSPDOEDC 480
Db 421 KSLIFCDNRNGSKFRYLLHDQVDFLTKNCSQLODLHKKIITQORVHOPHTLSPDOEDC 480
Qy 480 KSLIFCDNRNGSKFRYLLHDQVDFLTKNCSQLODLHKKIITQORVHOPHTLSPDOEDC 480
Db 480 KSLIFCDNRNGSKFRYLLHDQVDFLTKNCSQLODLHKKIITQORVHOPHTLSPDOEDC 480

```



```

QY 481 MYWYNFLAYHMASAKHKECALMFSLDWIKAKTELVGPAHLIHFVEYRHLDEKDCAV 540
Db 470 MYWYNFLAYHMASAKHKECALMFSLDWIKAKTELVGPAHLIHFVEYRHLDEKDCAV 529
QY 541 SENFOEFLSINGHLLGRPFNNIVQLGLCEPSETSEYVQAKLQAKQKQEVNDGMLYLEWINK 600
Db 530 SENFOEFLSINGHLLGRPFNNIVQLGLCEPSETSEYVQAKLQAKQKQEVNDGMLYLEWINK 589
QY 601 KNITNLRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFAETGEKLEIKAHDEV 660
Db 590 KNITNLRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFAETGEKLEIKAHDEV 649
QY 661 LCCAFSTDDRFATATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
Db 650 LCCAFSTDDRFATATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 709
QY 721 SDCFLKLDNLNKECRNTMFHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 780
Db 710 SDCFLKLDNLNKECRNTMFHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 769
QY 781 INVQOFFLNLEDPQEDMEVIVKCCSWADGARIWVAANKKIF----- 811
Db 770 INVQOFFLNLEDPQEDMEVIVKCCSWADGARIWVAANKKIF----- 811
QY 841 STIQYCDSPQNHVALVALSOYCVELNWTDSRSKVADCRGHLVSWHGVMSFPGSSFLTS 900
Db 812 ----- 812
QY 901 SDQTLRLWETKVKCNKSAVMLKQEVNDVFOENEVNLAVDHRLQLINGRTGQIDYLT 960
Db 847 SDQTLRLWETKVKCNKSAVMLKQEVNDVFOENEVNLAVDHRLQLINGRTGQIDYLT 906
QY 961 EAQVSCCCLSPHLOQYIARFENGDAIELELVNNRIFQSRFQHKKTWVHIOPTADEKTLIS 1020
Db 907 EAQVSCCCLSPHLOQYIARFENGDAIELELVNNRIFQSRFQHKKTWVHIOPTADEKTLIS 966
QY 1021 SSDDAEIQWNLQDKCIFLGHQETVKDFRLLKNSRLSSFDGTVKVNIIITGNKEKD 1080
Db 967 SSDDAEIQWNLQDKCIFLGHQETVKDFRLLKNSRLSSFDGTVKVNIIITGNKEKD 1026
QY 1081 FVCHQGTVLSCDISHDATKSSSTADKTKINSFDLLPLHELGHGNCVRCASFVDSST 1140
Db 1027 FVCHQGTVLSCDISHDATKSSSTADKTKINSFDLLPLHELGHGNCVRCASFVDSST 1086
QY 1141 LLATGDNGEIRIWNVNGELLHLICAPLSEGNATHGGWVTDLCFSPDGKMLISAGGYIK 1200
Db 1087 LLATGDNGEIRIWNVNGELLHLICAPLSEGNATHGGWVTDLCFSPDGKMLISAGGYIK 1146
QY 1201 WNNVVTGSSQFYNYNGNLKKIHVSPDKTYVYVNDNLGILYILOTLE 1248
Db 1147 WNNVVTGSSQFYNYNGNLKKIHVSPDKTYVYVNDNLGILYILOTLE 1194

```

```

RESULT 4
US-10-141-618-10
; Sequence 10, Application US/10141618
; Publication No. US20030165887A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: Methods For Determining the Prognosis
; OF INVENTION: For Cancer Patients Using Tucan
; FILE REFERENCE: P-LJ 5254
; CURRENT APPLICATION NUMBER: US/10/141,618
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/289,233
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: US 60/356,934
; PRIOR FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: US 09/388,221
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 10

```

```

; LENGTH: 1194
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-141-618-10

Query Match 94.9%; Score 6282; DB 14; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

QY 1 MDKARNCLLQHRLEAKEDIKTSYIMDHMISDGLFTTISEBEKVNEPTQOORAAMLIKMI 60
Db 1 MDKARNCLLQHRLEAKEDIKTSYIMDHMISDGLFTTISEBEKVNEPTQOORAAMLIKMI 60
QY 61 LKKNDSVSVFYNALLHEGYKDLAALLHDGIPVSSSGKDSVSGITSYVTVLCEGGVP 120
Db 61 LKKNDSVSVFYNALLHEGYKDLAALLHDGIPVSSSGKDSVSGITSYVTVLCEGGVP 109
QY 121 QRPVYVTRKLVNAIOOKLSKLGEPGWVTHMGAGCGKSVLAENVRDHSLLGECFPG 180
Db 110 QRPVYVTRKLVNAIOOKLSKLGEPGWVTHMGAGCGKSVLAENVRDHSLLGECFPG 169
QY 181 GVHWVSVCKQKSGELLMLKQNLCTRLDQDESFSQRLPLNIEAKORLRLILMRKXPRSL 240
Db 170 GVHWVSVCKQKSGELLMLKQNLCTRLDQDESFSQRLPLNIEAKORLRLILMRKXPRSL 229
QY 241 ILDDVMSWLVKAFDSQOQILLTTRDKSVTDSVMGPKYVVPVSSSLGKGLFLLSFVN 300
Db 230 ILDDVMSWLVKAFDSQOQILLTTRDKSVTDSVMGPKYVVPVSSSLGKGLFLLSFVN 289
QY 301 MKKADLPQEAHSIIKECGSPVLSLIGALLRDPNRMWYVYLKQKQPKRIRKSSSYD 360
Db 290 MKKADLPQEAHSIIKECGSPVLSLIGALLRDPNRMWYVYLKQKQPKRIRKSSSYD 349
QY 361 YEALDEAMSISVEMLRDIDKYVYTSLIQDKVPTKVLICILWDMETEVEDILQEFVN 420
Db 350 YEALDEAMSISVEMLRDIDKYVYTSLIQDKVPTKVLICILWDMETEVEDILQEFVN 409
QY 421 KSLFLCDNRGKSPRYVLDLQVDFLTENKCSQLODLHKLIITQFQYHOPHTLSPQEDC 480
Db 410 KSLFLCDNRGKSPRYVLDLQVDFLTENKCSQLODLHKLIITQFQYHOPHTLSPQEDC 469
QY 481 MYWYNFLAYHMASAKHKECALMFSLDWIKAKTELVGPAHLIHFVEYRHLDEKDCAV 540
Db 470 MYWYNFLAYHMASAKHKECALMFSLDWIKAKTELVGPAHLIHFVEYRHLDEKDCAV 529
QY 541 SENFOEFLSINGHLLGRQPPNNIVOLGLCEPSETSEYVQAKLQAKQEVNDGMLYLEWINK 600
Db 530 SENFOEFLSINGHLLGRQPPNNIVOLGLCEPSETSEYVQAKLQAKQEVNDGMLYLEWINK 589
QY 601 KNITNLRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFAETGEKLEIKAHDEV 660
Db 590 KNITNLRLVVRPHTDAVYHACFSEDGQRIASCADKTLQVFAETGEKLEIKAHDEV 649
QY 661 LCCAFSTDDRFATATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 720
Db 650 LCCAFSTDDRFATATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLATGS 709
QY 721 SDCFLKLDNLNKECRNTMFHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 780
Db 710 SDCFLKLDNLNKECRNTMFHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 769
QY 781 INVQOFFLNLEDPQEDMEVIVKCCSWADGARIWVAANKKIF----- 811
Db 770 INVQOFFLNLEDPQEDMEVIVKCCSWADGARIWVAANKKIF----- 811
QY 841 STIQYCDSPQNHVALVALSOYCVELNWTDSRSKVADCRGHLVSWHGVMSFPGSSFLTS 900
Db 812 ----- 812
QY 901 SDQTLRLWETKVKCNKSAVMLKQEVNDVFOENEVNLAVDHRLQLINGRTGQIDYLT 960
Db 847 SDQTLRLWETKVKCNKSAVMLKQEVNDVFOENEVNLAVDHRLQLINGRTGQIDYLT 906

```


QY 1096 DATKFSSTADKATKIMSDLLPLHELHNGCVRCSAFSDVSTLLATGDNDNGEIRIWN 1155
 DB 1047 DATKFSSTADKATKIMSDLLPLHELHNGCVRCSAFSDVSTLLATGDNDNGEIRIWN 1106
 QY 1156 VSNGLLHLCAPLSEGAATHGHWTDLCFSPDGKMLISAGYIKWNNVVTGSSOTFTY 1215
 DB 1107 VSNGLLHLCAPLSEGAATHGHWTDLCFSPDGKMLISAGYIKWNNVVTGSSOTFTY 1166
 QY 1216 NGTNLKKHVSDFKTYVVDNIGILYILOTLE 1248
 DB 1167 NGTNLKKHVSDFKTYVVDNIGILYILOTLE 1199

RESULT 6

US-10-112-944-272
 ; Sequence 272, Application US/10112944
 ; Publication No. US20040048249A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Gezhi
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Ghosh, Malabika
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wang, Zhiwei

TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
 ; TITLE OF INVENTION: Secreted Polypeptides
 ; FILE REFERENCE: 805A
 ; CURRENT APPLICATION NUMBER: US/10/112, 944
 ; PRIOR FILING DATE: 2002-03-28
 ; PRIOR FILING DATE: US 09/488, 725
 ; PRIOR FILING DATE: 2000-01-21
 ; PRIOR APPLICATION NUMBER: US 09/491, 404
 ; PRIOR FILING DATE: 2000-01-25
 ; PRIOR APPLICATION NUMBER: US 09/496, 914
 ; PRIOR FILING DATE: 2000-02-03
 ; PRIOR APPLICATION NUMBER: US 09/515, 126
 ; PRIOR FILING DATE: 2000-02-28
 ; PRIOR APPLICATION NUMBER: US 09/519, 705
 ; PRIOR FILING DATE: 2000-03-07
 ; PRIOR APPLICATION NUMBER: US 09/540, 217
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: US 09/552, 929
 ; PRIOR FILING DATE: 2000-04-18
 ; PRIOR APPLICATION NUMBER: US 09/577, 408
 ; PRIOR FILING DATE: 2000-05-18
 ; NUMBER OF SEQ ID NOS: 924
 ; SOFTWARE: pt_FL_genes Version 5.0
 ; SEQ ID NO 272
 ; LENGTH: 1526
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

Query Match 89.6%; Score 5933.5; DB 15; Length 1526;
 Best Local Similarity 89.1%; Pred. No. 0;
 Matches 1141; Conservative 0; Mismatches 0; Indels 139; Gaps 5;
 QY 14 EALEKDITKSYMDHMSDGLTISBEEKVRNEPTQOQRAAMLKMLKKDNDSYVSFY 73
 DB 341 EALEKDITKSYMDHMSDGLTISBEEKVRNEPTQOQRAAMLKMLKKDNDSYVSFY 400
 QY 74 ALHLEGYKDLAALLHDGIPVWSSSKGDSVSGITSYRTVLCEGGVQRPVVFVTRKKLV 133
 DB 401 ALHLEGYKDLAALLHDGIPVWSSSS-----VTVLCEGGVQRPVVFVTRKKLV 449
 QY 134 NAIQOKLSKLGEPGWTTIHGMAGCKSVLAEEAVRDHSLLEGCFPGGVHWSVGKQDKS 193

DB 450 NAIQOKLSKLGEPGWTTIHGMAGCKSVLAEEAVRDHSLLEGCFPGGVHWSVGKQDKS 509
 QY 194 GLMKLQNLCTRLDQDESFSORPLNIEEAKDRILMLRKHPRSLILDDVWDSVWLKA 253
 DB 510 GLMKLQNLCTRLDQDESFSORPLNIEEAKDRILMLRKHPRSLILDDVWDSVWLKA 569
 QY 254 FDSQCOILLTRKSVTDSVMGPKYVVPVSSLSGKEGLETLISLVNMMKADLPQAHSI 313
 DB 570 FDSQCOILLTRKSVTDSVMGPKYVVPVSSLSGKEGLETLISLVNMMKADLPQAHSI 629
 QY 314 IKECKGSPVLSIGALLRDFNRWEYILKOLNOKPRTRKSSSYDEALDEMSISVE 373
 DB 630 IKECK-----VSLIGALLRDFNRWEYILKOLNOKPRTRKSSSYDEALDEMSISVE 684
 QY 374 MLREDIKDYTTDLSILOKDVKPTKVLCLMDMETEVEDILOEFVNKSLLCDRNGKSF 433
 DB 685 MLREDIKDYTTDLSILOKDVKPTK----- 709
 QY 434 RYLLHDLQVDFLTERKNCSQLDLHKKIITQFYHQPHTLSPQEDCMWYNFLAYHMAS 493
 DB 710 -----DLHKKIITQFYHQPHTLSPQEDCMWYNFLAYHMAS 748
 QY 494 AKMHKELCALMFSLDWIKARTELVGPAHLIHEFVEYRHLDEKDCAVSENFOFFLSLNGH 553
 DB 749 AKMHKELCALMFSLDWIKARTELVGPAHLIHEFVEYRHLDEK----- 791
 QY 554 LLGRQPPFNIVOLGLCEPETSEVYQOAKLOAKQKQVNDGMLYLEWINKNITNLSRLVVRP 613
 DB 792 -----PPFNIVOLGLCEPETSEVYQOAKLOAKQKQVNDGMLYLEWINKNITNLSRLVVRP 846
 QY 614 HTDAVYHACFSEOGRIASCGADKTLQ----- 640
 DB 847 HTDAVYHACFSEOGRIASCGADKTLQLOARASGSTRQAGPLPPGAARAAQAAVADSZIP 906
 QY 641 -----VFKAETGEKLEIKAHDEVILCASTDDRIATCSVDKVKIWNST 688
 DB 907 VLGLGEGPALVFKAETGEKLEIKAHDEVILCASTDDRIATCSVDKVKIWNST 966
 QY 689 GELVHTYDEHSEQVNCCHFTNSHHLLATGSSDCFLKLDLNQKCRNTMFGHTNSVNH 748
 DB 967 GELVHTYDEHSEQVNCCHFTNSHHLLATGSSDCFLKLDLNQKCRNTMFGHTNSVNH 1026
 QY 749 CRFSPDDKLLASCSADGTLKLDATSAHERKSINVKOFFLNLEDPOEDMEVIVKCCWSA 808
 DB 1027 CRFSPDDKLLASCSADGTLKLDATSAHERKSINVKOFFLNLEDPOEDMEVIVKCCWSA 1086
 QY 809 DGARIMVAAKNKIFLFDIHTSGLLGEIHTGHSTIQYCDFSPQNLAVVALSQYCVELWN 868
 DB 1087 DGARIMVAAKNKIFLFDIHTSGLLGEIHTGHSTIQYCDFSPQNLAVVALSQYCVELWN 1146
 QY 869 TDSRSKVADCRGHLVWGVWVSPGSSFLTSSDDQIRLWETKVKCKNSAVMLKQEVVDV 928
 DB 1147 TDSRSKVADCRGHLVWGVWVSPGSSFLTSSDDQIRLWETKVKCKNSAVMLKQEVVDV 1206
 QY 929 VFOENEVWVLAVDHIRRLQIINGRTGQIDYLTAEQVSCCLSPHLOYIAFGDENGAIETL 988
 DB 1207 VFOENEVWVLAVDHIRRLQIINGRTGQIDYLTAEQVSCCLSPHLOYIAFGDENGAIETL 1266
 QY 989 ELVNNRIFQSRFOHKTVWHIOFTADEKTLISSDDAEIOVWNNQDKCIFLGHQETVK 1048
 DB 1267 ELVNNRIFQSRFOHKTVWHIOFTADEKTLISSDDAEIOVWNNQDKCIFLGHQETVK 1326
 QY 1049 DFRLLKNSRLLSWFGTVKWNIIITGNKEKDFVCHOGTVLSCDI SHDATKFSSTADKT 1108
 DB 1327 DFRLLKNSRLLSWFGTVKWNIIITGNKEKDFVCHOGTVLSCDI SHDATKFSSTADKT 1386
 QY 1109 AKIWSFDLLPLHELHNGCVRCSAFSDVSTLLATGDNDNGEIRIWNVNGELLHLCAPL 1168
 DB 1387 AKIWSFDLLPLHELHNGCVRCSAFSDVSTLLATGDNDNGEIRIWNVNGELLHLCAPL 1446
 QY 1169 SEEGAATHGCVTDLCFSPDGKMLISAGYIKWNNVVTGSSOTFTYNTGNTLNKKHVSPD 1228

Db 1447 SEEGNATGGWTDLCPSDGMKLISAGYIKWNNVVTGESSQFYTTNGTNLKKIHVSPD 1506
Qy 1229 FKTYVTVDNLGLYILQTL 1248
Db 1507 FKTYVTVDNLGLYILQTL 1526

RESULT 7

US-10-014-269-31
; Sequence 31, Application US/10014269
; Publication No. US20020127673A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Ogura, Yasunori
; TITLE OF INVENTION: NOD2 Nucleic Acids and Proteins
; FILE REFERENCE: UM-06645
; CURRENT APPLICATION NUMBER: US/10/014,269
; PRIORITY FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 31
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-014-269-31

Query Match 24.0%; Score 1589; DB 13; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.8e-118;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 123 PVFVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCGKSVLAEEAVRDHSLLEGCFPGV 182
Db 1 PVFVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCGKSVLAEEAVRDHSLLEGCFPGV 60
Qy 183 HWSVGKDKSGLMKLQNLCTRLDQDSFSQRLPLNTEEAKDRILMLRKHPRSLIL 242
Db 61 HWSVGKDKSGLMKLQNLCTRLDQDSFSQRLPLNTEEAKDRILMLRKHPRSLIL 120
Qy 243 DDVWDSWVLKAFDSQCQILLTRDKSVTDSVMGPKYVVPVVESSLGKEKGLSILFVNKM 302
Db 121 DDVWDSWVLKAFDSQCQILLTRDKSVTDSVMGPKYVVPVVESSLGKEKGLSILFVNKM 180
Qy 303 KADLPEQAHSIIKECKGSPVLSLIGALLRDPFNNRWEYLLKQLQNKQPKRIRKSSSYDYE 362
Db 181 KADLPEQAHSIIKECKGSPVLSLIGALLRDPFNNRWEYLLKQLQNKQPKRIRKSSSYDYE 240
Qy 363 ALDEAMSI SVMELREDIKDYTTDLSILQKQVKTPLVLCILWDMETEVEEDILQEFVNKS 422
Db 241 ALDEAMSI SVMELREDIKDYTTDLSILQKQVKTPLVLCILWDMETEVEEDILQEFVNKS 300
Qy 423 LLFCDRNG 430
Db 301 LLFCDRNG 308

RESULT 8

US-10-002-974-31
; Sequence 31, Application US/10002974
; Publication No. US20020197616A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Ogura, Yasunori
; APPLICANT: Cho, Judy
; APPLICANT: Nicolae, Dan L
; APPLICANT: Bonen, Denise
; TITLE OF INVENTION: NOD2 Nucleic Acids and Proteins
; FILE REFERENCE: UM-06646
; CURRENT APPLICATION NUMBER: US/10/002,974
; PRIORITY FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: Patent in version 3.1

; SEQ ID NO 31
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-974-31

Query Match 24.0%; Score 1589; DB 13; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.8e-118;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 123 PVFVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCGKSVLAEEAVRDHSLLEGCFPGV 182
Db 1 PVFVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCGKSVLAEEAVRDHSLLEGCFPGV 60
Qy 183 HWSVGKDKSGLMKLQNLCTRLDQDSFSQRLPLNTEEAKDRILMLRKHPRSLIL 242
Db 61 HWSVGKDKSGLMKLQNLCTRLDQDSFSQRLPLNTEEAKDRILMLRKHPRSLIL 120
Qy 243 DDVWDSWVLKAFDSQCQILLTRDKSVTDSVMGPKYVVPVVESSLGKEKGLSILFVNKM 302
Db 121 DDVWDSWVLKAFDSQCQILLTRDKSVTDSVMGPKYVVPVVESSLGKEKGLSILFVNKM 180
Qy 303 KADLPEQAHSIIKECKGSPVLSLIGALLRDPFNNRWEYLLKQLQNKQPKRIRKSSSYDYE 362
Db 181 KADLPEQAHSIIKECKGSPVLSLIGALLRDPFNNRWEYLLKQLQNKQPKRIRKSSSYDYE 240
Qy 363 ALDEAMSI SVMELREDIKDYTTDLSILQKQVKTPLVLCILWDMETEVEEDILQEFVNKS 422
Db 241 ALDEAMSI SVMELREDIKDYTTDLSILQKQVKTPLVLCILWDMETEVEEDILQEFVNKS 300
Qy 423 LLFCDRNG 430
Db 301 LLFCDRNG 308

RESULT 9

US-10-314-506-31
; Sequence 31, Application US/10314506
; Publication No. US20030175762A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Ogura, Yasunori
; TITLE OF INVENTION: Modulators of NOD2 Signaling
; FILE REFERENCE: UM-06984
; CURRENT APPLICATION NUMBER: US/10/314,506
; PRIORITY FILING DATE: 2002-12-09
; PRIORITY FILING DATE: 2001-10-26
; PRIORITY FILING DATE: 60/244,289
; PRIORITY FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 31
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-314-506-31

Query Match 24.0%; Score 1589; DB 14; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.8e-118;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 123 PVFVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCGKSVLAEEAVRDHSLLEGCFPGV 182
Db 1 PVFVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCGKSVLAEEAVRDHSLLEGCFPGV 60
Qy 183 HWSVGKDKSGLMKLQNLCTRLDQDSFSQRLPLNTEEAKDRILMLRKHPRSLIL 242
Db 61 HWSVGKDKSGLMKLQNLCTRLDQDSFSQRLPLNTEEAKDRILMLRKHPRSLIL 120
Qy 243 DDVWDSWVLKAFDSQCQILLTRDKSVTDSVMGPKYVVPVVESSLGKEKGLSILFVNKM 302
Db 301 LLFCDRNG 308

Db 121 DVWDSWVLKAFDSQOILLTTRDKSVTDSVMGPKVVPVSSLGKEKGLLEILSLFVNNK 180
 Qy 303 KADLPQAHSIKECKGSPVLSIGALLRDPFNNRWEYLLKOLQNKQPKRIRKSSSYDYE 362
 Db 181 KADLPQAHSIKECKGSPVLSIGALLRDPFNNRWEYLLKOLQNKQPKRIRKSSSYDYE 240
 Qy 363 ALDEAMSIIVMLREDIKDYITDLSILQKDVKVPYTKVLKILWDMETBEVEDILQEFVNS 422
 Db 241 ALDEAMSIIVMLREDIKDYITDLSILQKDVKVPYTKVLKILWDMETBEVEDILQEFVNS 300
 Qy 423 LLFCDRNG 430
 Db 301 LLFCDRNG 308

RESULT 10
 US-10-001-254-29
 ; Sequence 29, Application US/10001254
 ; Publication No. US20030049702A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, John C.
 ; APPLICANT: Godzik, Adam
 ; APPLICANT: Pawlowski, Krzysztof
 ; APPLICANT: Fiorentino, Loredana
 ; APPLICANT: Lee, Sug Hyung
 ; APPLICANT: Roth, Wilfried
 ; APPLICANT: Stenner-Liewen, Frank
 ; TITLE OF INVENTION: No. US20030049702A1el Death Domain Proteins
 ; FILE REFERENCE: P-LJ 5037
 ; CURRENT APPLICATION NUMBER: US/10/001,254
 ; CURRENT FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: 60/301,889
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: 09/715,893
 ; PRIOR FILING DATE: 2000-11-17
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 29
 ; LENGTH: 142
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; US-10-001-254-29

Query Match 11.3%; Score 748; DB 14; Length 142;
 Best Local Similarity 100.0%; Pred. No. 1.4e-51;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 142 KLKGEFGWTHIGMAGCKSVLAEEAVRDHSLLEGCFGGVHWVSVGKQKSGLLMKLQN 201
 Db 1 KLKGEFGWTHIGMAGCKSVLAEEAVRDHSLLEGCFGGVHWVSVGKQKSGLLMKLQN 60
 Qy 202 LCTRLDQDESORPLNIEAKDRILMLKHPRSLLILDVWDSWVLKAFDSQOILL 261
 Db 61 LCTRLDQDESORPLNIEAKDRILMLKHPRSLLILDVWDSWVLKAFDSQOILL 120
 Qy 262 LTRDKSVTDSVMGPKVVPVE 283
 Db 121 LTRDKSVTDSVMGPKVVPVE 142

RESULT 11
 US-10-369-493-12913
 ; Sequence 12913, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 12913
 ; LENGTH: 529
 ; TYPE: PRT
 ; ORGANISM: Aspergillus nidulans
 ; US-10-369-493-12913
 Query Match 10.7%; Score 711; DB 14; Length 529;
 Best Local Similarity 29.0%; Pred. No. 1e-47;
 Matches 183; Conservative 93; Mismatches 242; Indels 112; Gaps 11;
 Qy 614 HTDAVTHACFSEDSQRTASCGADKTLQVFAETGKLEIKAHEDVLCARSTDRFTA 673
 Db 4 HLGPVESVVFSPDGKQLVSGSYDDTVKIWDPATGELLQNDGHSVTESLAFSPDGKLLA 63
 Qy 674 TCSVDKVKIWNMTGELVHTYDEHSEOVNCHFTNSHHLLATGSSDCLKLDLQNK 733
 Db 64 SGSYDNTIDLWDSPYTGELQTFEGHPHSIWSVAPADGKE--LASASDDSTIKIWDLATG 121
 Qy 734 ECRNTMGHTNSVNHCRFSPDDKLLASCADCTKLWDATSANERKSINVQFFLNEDP 793
 Db 122 ELQQLDHSQSRSVAFSPDGKLLASSLSDSTIKVWNPATGELQQL----- 169
 Qy 794 QDMEVIVKCCSWASADGARIMVAA--KNKIFLFDIHTSGLLGEIHTGHSTIQYCDSPON 852
 Db 170 --EGRSGWVKVAFSPDGKLLASGSEKNTVKLWNPATGELIQTLE--GHSOSVRSVAFSPDG 227
 Qy 853 HLAVALSQVCVELMNTDSRSKVADCRGHLNVHGVMPDGSFLTSSDDOTIRLWETK 912
 Db 228 KQLASSSDTTIKLWNTTGGELQQTQFKGDLWIRAVAFSPDGKHLVSGSDDNITIKLW--- 284
 Qy 913 KYCKNSAVMLKQEVVDVFOENEVMVLAVDHIRRLQINGRTGQIDYLTEAQVSCCCLSPH 972
 Db 285 ----- 284
 Qy 973 LQYIAFGDENGAIETLELVNRIQSRFHKKTVWHIQTADKTLISSDDAEIQVWN- 1031
 Db 285 -----DLATSELOQSLSDHSRSVHAFSPDDKQLASSLSDSTIKLWDS 328
 Qy 1032 --WOLDKCIFLRGHQTVDKFRLLKNSRL--SWFSDGTGVKVNIIITGNKEKDFVCHQTV 1088
 Db 329 ATGELQRT--LEHSGQVRSVTFSPDGKLLASNSYDGTIKLWNPATGELQQTLTGRSDWV 386
 Qy 1089 LSCDISHDATKFSSTADKTAKIWSFDLPLHLRGHNGVRCSAFSDVSTLLATGDDN 1148
 Db 387 DSAVSPDGKQLASGYDSTIKLWDSATGELLQTLLEGSDRIQSVVFPDGKLLASGPTN 446
 Qy 1149 GEIRIWNVSGELLHLCAPLSEGAATHGHWVTLCSFDGKMLISA--GGYIKWNVVT 1206
 Db 447 KPAKFWDPATGELQIF-----EG--HKKWVESVAFSPDGKLLASSYGETIKLWDPT 498
 Qy 1207 GRSQQTFTYTGNTLKKIHVSDFKTYVTVD 1236
 Db 499 GELLQTLNDPDESAGSVAFSPDGNRLASVD 528

RESULT 12
 US-10-369-493-20185
 ; Sequence 20185, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 20185
 ; LENGTH: 610
 ; TYPE: PRT
 ; ORGANISM: No. US20030233675altoc punctiforme
 US-10-369-493-20185

Query Match	10.4%	Score 690.5	DB 14	Length 610
Best Local Similarity	28.7%	Pred. No. 5.7e-46	Indels 85	Gaps 14
Matches 187	Conservative 106	Mismatches 273	Indels 85	Gaps 14

Qy	614	HTDAVYHACFSEDGQRIASCGADKTLQVFAETGEKLEIKAHEDVILCCAFSTDDREIA	673
Db	5	HAKVQGISFSPDGKMLASADNDNTVKLWDTTGTKEIKTLTGHTNSVLGISFSPDGKMLA	64
Qy	674	TCSVDKVKVWNSMTGELVHTVDEHSEQVNCCHFTNSSHLLLATGSSDCFLKWLNDLQK	733
Db	65	SASDNTVKLWDTTGTKEIKTLTGHTNSVLGISF--SPDGKMLASASADNTVKLWDTTGT	122
Qy	734	ECRNTMFGHTNSVNHCRSPDGKMLASADNTVKLWDTTGTKEIKTLTGHTNSVLGISF	793
Db	123	KEIKTLTGHTNSVFGISFSPDGKMLASADNTVKLWDTTGTKEIKTLTGHTNSVLGISF	170
Qy	794	QEDMEVVKCCSWSDAGARIMAAKNKIFLFDIHTSGLLGELIHTGHSTIOYCDSPQNH	853
Db	171	-----TGRNSVFGISFSPDGK 187	
Qy	854	LAVVALSYCVELWNTDGRSKVADCRHLVWGHVWVSPDGSSFLTSSDDOTIRLWETK-	912
Db	188	MLASASFDNTVKLWDTTGTKEIKTLTGHTNSVNDISFSPDGKMLASADNTVKLWDTT	247
Qy	913	-KVCNSAVMLKQEVVVQFQNEVWVAVDHRRLQINGRTG-QIDVLT--EAQVSCC 968	
Db	248	GKEIKTLTGHTNSVNDISFSPDGKMLASASFDNTVKLWDTTGTKEIKTLTGHTNSVNDIS	307
Qy	969	LSPLHQLYAFDGENGAILELVNRRIFQSFQHKKTVMHIOFTADEKTLISSSDABEQ 1028	
Db	308	FSPDGKMLASAGDNTVKLWDTTGTKEIKTLTGHTNSVNDISFSPDGKMLASAGDNTVK 367	
Qy	1029	VWNLQDKCI-FLRGHQTVDKFLRLKNSRL--SWSPDGTGVKWNIIITGNKEKDFVCHQG 1086	
Db	368	LWDTTGTKEIKTLTGHTNSVNGISFSPDGKMLASAGDKTVKLDWDTTGTKEIKTLTGHTN 427	
Qy	1087	TVLSCDISHDATKFSSTADKTAKIWSFDLLPLHLRGHNGCVRCASFVSDTLLATGD 1146	
Db	428	SVNGISFSPDGKMLASAGDKTVKLDWDTTGTKEIKTLTGHTNSVNGISFSPDGKMLASAS 487	
Qy	1147	DNGEIRIWN-VSNGELLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAG--GYIKWN 1203	
Db	488	SDNTVKLWDTTGTKEIKTLTGHTNSVNGISFSPDGKMLASADNTVKLWDTTGTKEIKTL 539	
Qy	1204	VVTGESSQTFYNTGN-LKKIHVSPDK--TYVTVDN-----LGLYLQ 1245	
Db	540	TTTGTKEIKTL-TGHTNWYVIGISFSPDGKMLASADNTVKLWDLDFLLQ 589	

RESULT 13

US-10-369-493-20242
 ; Sequence 20242, Application US/10369493
 ; Publication No. US20030233675altoc punctiforme
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 20242
 ; LENGTH: 1005
 ; TYPE: PRT
 ; ORGANISM: No. US20030233675altoc punctiforme
 US-10-369-493-20242

Query Match	10.4%	Score 688	DB 14	Length 1005
Best Local Similarity	24.6%	Pred. No. 2e-45	Indels 140	Gaps 21
Matches 212	Conservative 163	Mismatches 348	Indels 140	Gaps 21

Qy	407	ETEVEDILOFVFNKSLIFCDNRNGKSPRYYL-HDLQVDFLTKXNSQIQDLHKHKKIITQFQ 465
Db	239	EINKLDLILEIFVQSLVLLPENADRYQLVHDYIAAFIRQQQEPKL----KQVMAELE 294
Qy	466	RYHQPHLTSPQEDCMYWNFLAYHMASAKMKELCALMFLSDMTAKTELVGPAHLIHE 525
Db	295	KERKERKLS-----EAKLNSFLKRALFST--VAAGLGAGLAAATFQ 334
Qy	526	FVEYRHILDEKCAVSENFQEFLSINGH-----LLGROPPFNIVOL-----GLCEPE 572
Db	335	W-----AVEANVNOISAINNSSEAFALSEKYPDALIAALKAGSKLKHTLWAOH 382
Qy	573	TSEVYQAKLQAKQEVNGMLYLEWINKNTNLISRLVVRPHTDAVYHACFSEDCQRIAS 632
Db	383	RSIDILMTVTLQAV-----YLKPKCKENRALEVTNTEGHSDWVSSVAVSPNGYQLAS 437
Qy	633	CGAKTLQVKAETGEKLEIKAHEDVLCVCAFSTDDRFIATCVSDKVKKIVWNSMTGBLV 692
Db	438	ASADTKTIKWDVSSGQLKTLTGHSDRISIRIAYSNGQQLVVSADKTIKIWDVSSGKLL 497
Qy	693	HTYDEHSEQVNCCHFTNSSHLLLATGSSDCFLKWLNDLQNEKCRNTMFGHTNSVNHCRFS 752
Db	498	KTLTGHTSAVSSVAVNPNQO--LASASDDNTIKIWDISSGKLLKTLTGHSVSVNSVAVN 555
Qy	753	PDDKLLASCADGTLKWLADATSAERKSINVKQFPLENLEDPQEDMEVIVKCCSWADGAR 812
Db	556	PNQQLASASNDKTIKWDINSKLLKSLTGHSSEVN-----SVAVSPNGQQ 602
Qy	813	IMVAA-KNKIPLFDIHTSGELGEIHTGHSTIOYCDSPQNHAVVALSYCVELWNTDS 871
Db	603	LASASFDNTIKIWDI--SSGKLLKTLTGHSNVVSVAVSPNGQHLASASADKTIKIWDVSS 661
Qy	872	RSKVADCRHLVWGHVWVSPDGSSFLTSSDDOTIRLWETKVKCKNSAVMLKQEVVDVFPQ 931
Db	662	GKPLKSLAGHSNVVSVAVSPNGQQLASASDDKTIKWD----- 700
Qy	932	ENEVWVLAVDHIRRLQLINGRTGQIDVLT--AOVSCCLSPHLQYIAFGDENGAIETLE 989
Db	701	-----ISNGK--PLEMTDHSDRVSVVSPNGQHLASPSYDKTIKILWN 742
Qy	990	LWNRRIFQSRFQHKKTVMHIOFTADEKTLISSSDAEIQVWNLQDKCI-FLRGHQTVK 1048
Db	743	VSSGKLLKTLTGHSSEVNSVAVSPNGQQLASASWDKTIKWDVNSGKPLKTLTGHSVNV 802
Qy	1049	DFRLKLN-SRLLSWSFDGTGVKWNIIITGNKEKDFVCHQGTVLSCDISHDATKFSSTADK 1107
Db	803	SVAVSPNGQQLASASFDNTIKIWDVSSGKLLKTLTGHSNAVSVAVSPNGQQLASASLDN 862
Qy	1108	TAKIWSFDLLPLHLRGHNGCVRCASFVSDTLLATGDDNGEIRIWNVNSNGELLHLCAP 1167
Db	863	TIKLDVSSAKLLKTLTGHSNAVSVAVSPNGQQLASASDDNTIKIWDVSSGKLL----- 917
Qy	1168	LSEGAATHGGWVTDLCFSPDGKMLISAG--GYIKWNVVTGESSQ-----TF 1213
Db	918	---KSLSGHSNAVSVIAYSNGQQLASASADNTIKIWDVSSGKLLKSLSGHSDWVMRTY 974
Qy	1214	YTNGTNLKKIHVSPDKTYVTVD 1236

Db 975 NPNGQQLASAVD---KTIILWD 994

RESULT 14

US-10-369-493-19053
; Sequence 19053, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19053
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Anabaena PCC7120
US-10-369-493-19053

Query Match 10.2%; Score 674; DB 14; Length 1140;
Best Local Similarity 24.2%; Pred. No. 3.3e-44;
Matches 293; Conservative 223; Mismatches 469; Indels 228; Gaps 57;

Qy 126 FVTRKKLVNAIQKLSKLGKSPGFWITHMGAGCKGSV-----LAABVRDHSLLGCGFCG 180
Db 21 FYGRETLT---QQLIADLANK-SFMAYVGASGSGKSVVQAGLIAQLRRGQOL-----PG 71
Qy 181 GVHWVSGKQDGSGLMKLQN-LCTRLDQDESESRQLPLN---IBEAKDLRLILMLRKHP 236
Db 72 SQEMWMSFRPEYPLVLSLCHLVDSTGEKAYQOMQLEGNLYQGAQGFVHMLHRRRP 131
Qy 237 RSLIILDDVDSWLVKAFDSQCQILLT-----TRDK-----SVTDSVMGPKYVVPVES 284
Db 132 MVVLVDQEEELFTLAASEDRQRFLDTVLGALELSPDKFKLIVTLRADFIAPCLEIPTLA 191
Qy 285 SLGKEKGL-----EILSLFN-MKKADL---PEQAHSIIKECKGSLPLVSLIGALL 331
Db 192 KLLQOSSILLPCLTQEBYRRIIHPAEKVGTLVDPELVLLQELHNSPGDLPLEFVL 251
Qy 332 RDPFNWEY-----YKQLOKQKFRIRKSSSYDYALD-EAMS-----IS 371
Db 252 EQL---WEYRDKGVITLQAYQYLGKIG-----ALEKAAQGVYDLDESEAEQCTRWIFLS 304
Qy 372 VEMLRD-----IKDYTDLSILOKOVKPTKVLKILMDMETEE-----410
Db 305 LTOLGEGTETRRRVLKSELIVKYY--PVALVERTLQVLTAKLVVWNGWESGRGKQ 362
Qy 411 -----VEDILO-----EFVNSKLLFCDRNGKGFYYLHDLQVDFLTKNCSQLQDLH 457
Db 363 AGRGCGENILTTPSVTIEVAHEVII---RYWSTLRWWL-----BENSRILRS-H 409
Qy 458 KKIITQRYHQPHLTSPD-----QEDCMYW--YNFL--AYHMASAKM 496
Db 410 RO-IGESAAWQNNQPPDFLOGLVRLAEAEIYLYNIDELSDVDYQHFIEACLHERRRQ 468
Qy 497 HKELCALMFLSDWIKAKTEL---VGPALHIEFVEYRHILDEKDC-AVSENF-----QSF 547
Db 469 NQESRLRQARVASIISTLGLTAFGLAVFAYQOTQNAQLKEIQALNSLSENFLSHKQL 528
Qy 548 LSLNGHLLGRQFPFNIYVGLG---CEPETSEVYQAKQAKQOEVNDGMLYLEWINKNLT 604
Db 529 EALMTSVQAGKEVQNI-SLGIPPTDTQTATILQNAVYNTQER--NRLLNHW-----VT 580
Qy 605 NLSRLVVRPHTDVAVHACFESDQRTASCAGDKTLQVFAETGKLEIKAHDEVLCCA 664

Db 581 SVS-----YSPDGEVIASGVNDTILHWRD-GKLLTTLTGNDGVNSVS 624
Qy 665 FSTDRFIATCSDVKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGSSDCF 724
Db 625 FSPDGEIILASADSTIKLW-QRNGQLITTLKGHDQGVKSVSPNGE--IIASGSDHT 681
Qy 725 LKWLNDLNOKECRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKSKINVK 784
Db 682 INLWSRAG-LLLSLNGHSQGVNSIKFSPGEGDTIASASDDGTIRLW---SLDGRPLITI- 736
Qy 785 QPFLNLEDPQEDMEVIVKCCSWMSADGARIMVA-AKNKIFLFDIHTSGLLGEIHTGHSSTI 843
Db 737 -----PSHTKQVL--AVTFSPDQOTIVSAGADNTVKLWS--RNGTLTLTLEHNEAV 784
Qy 844 QXCDFSPQNHVLVALSOYCVELWNTDSRKSVDRCRGLHSMVHGWMFSPDGSSFLTSSDD 903
Db 785 MQVIFSPDGRLIATASADTKITLWSRDG-NILGTFAGHNHEVNSLSFSPDGNILASGDD 843
Qy 904 QTRILWETKVKCKNSAVMLKQEVVVFOENE---VMVLAVDHIHRELQINGRTGQIDYLT 960
Db 844 NTVRLMTVNRITLPKTFYGHKGSVYRFSNDOKKITSLSDTMKTWSLDGKLLQTLSSP 903
Qy 961 EAQVSCCLSPHLOYAFDENGATIBILELVNR---IFQSRFOHKTKTVWHIOFTRADEKT 1017
Db 904 LPDVTSISFTPDNKIVALASPDTHLY-----NQGGLLRSLPGHNHWTLSFSFNKOI 959
Qy 1018 LISSDDDAEIQVM--NWQDKCIFLRGHOETVKDFRLLKNR--LLSWSPDGTVKVWNIIIT 1074
Db 960 LASGSADTKIKLMSVNGRLKT--LLGHNGWVTDIKFSADGNIVSASADTKIKMS-LD 1016
Qy 1075 GNKEKDFVCHQGTVLSCDISHDATFSSSTADTKAKIWSFDLLPLHLRGNHNGVCVCSA 1134
Db 1017 GRLIRTLQHSASVWSVNLSPDQTLASTQDETIKLWNLNGEL-IYTLRGHSDVVYNLS 1075
Qy 1135 FSVSTLLATGDDNGEIRIWNVNGELLHLCAPLSEGAATHGCGWVTDLCFSPDGKMLIS 1194
Db 1076 FSPDKTITASDDGTIKLWNVNGTLK-----TFQG---HRGGVRSVSFSPDGKILAS 1127
Qy 1195 AGGY---IKWNVN 1204
Db 1128 -GGHDTTVKVNWL 1139

RESULT 15

US-10-369-493-19046
; Sequence 19046, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19046
; LENGTH: 1136
; TYPE: PRT
; ORGANISM: Anabaena PCC7120
US-10-369-493-19046

Query Match 10.2%; Score 672; DB 14; Length 1136;
Best Local Similarity 23.2%; Pred. No. 4.7e-44;
Matches 281; Conservative 196; Mismatches 490; Indels 244; Gaps 41;
Qy 95 SSSGKDSVSGISYRVTLCRGGVPOPPVTVTRKLVNAIQKLSKLGKSPGFWITHG 154

Db 46 ASGSGKSS-----VQAGLIPQ-----LROGKHIPNSBQWGIKTIR--PGVNPLEA 89
 Qy 155 MAGCKSVLAABAVRDHSLLEGCF-----PGGVHVSVKQDKSGLLM-----KLQNLCTRL 206
 Db 90 LAR-----KLGEWGETHLLIEGMLHQGVFVYMLRSLPQGVTVLVDIQFBELEFLTAPT 144
 Qy 207 DQDESFSQRLPLNIBEAQRDLRLMLRKHPRLSLIILDDVWDSWVLKAFDSQCQIILLTRD 266
 Db 145 DR-EUFLLELLGAVQAGDRFKLI-----TLRA----- 172
 Qy 267 KSVTDSVMGPKYVVPVVESSLGKEKLEILSLFVNMMKADLPQAHSLIIECKE--GSLPLV 324
 Db 173 -----DFIAPCLEVPALAEALQVAVSLV-PPKLSLDYRRVRLHPAQOVLQVEA 221
 Qy 325 SLIGALLR-----DFPNRWYYLKLQNKOKFKIRKSSSYD----- 360
 Db 222 ELVEVLLRELNGSVGDLPL-LLEFVLEQLWQORTAGKLTLSQYEQELGGIKGALERSCOGV 280
 Qy 361 YEALDEAMS-----ISVEMLREDIKO-----YVTDL-----SILQKOVKVPKVL 401
 Db 281 YESLPPQLQSCAKWIFLSLTQLGEGTEPTRRIYKSDLIIVKYPAGLVEQTLNVLNNAKL 340
 Qy 402 ILWDMETEVEDILOEFVNKSLFLCDRNKSPRYLHDLQVDFLTEKNCSQLQDLHKKII 461
 Db 341 VVINLEAE-----IEAQKSPSPPTPNLSTPFVTVEVAHEILIRHWSYL 384
 Qy 462 TOFORYHQPHTLSPDQ--EDCMYVY-----NFLAYHMASAKMKELCALMFSLDWIRAK 513
 Db 385 RWMLEENRDLRKQRIINHACQLWQSGKQADFL--LQGARLAEADIIYVWTDDELGAD 441
 Qy 514 T-ELVGP--AHLIHEFVEYRHILDEKCAV-----SENQFEL 548
 Db 442 VQEFIGACLAERKHQOLQAKRNLQKQARAVALSVLGIASVSGGLAYWQGREAQFREIA 501
 Qy 549 SLN-----GHLGRQPPFNVLQGLCEPETSEVVQAKLQAKQOEVDNGMLYLEWINKNIT 604
 Db 502 ALNSSQANLLSHQQLAALI-----ASLKAQOVNVHVIAPNNLKLATVT 546
 Qy 605 NLSRLV-----VRPHTDAYYHACFSEDQGIASCGADKTLQVFKAEKGKLEIKAH 656
 Db 547 TLQALFEMOERNLEGHKGQVISISIRDGGTIASGSLDKTIKLMSRD-GRLPFTLNGH 605
 Qy 657 EDEVLCCAFSTDDRIATCSVDKVKIWNMTGELVHYVDEHSEOVNCHFTNSSHLLL 716
 Db 606 EDVAVSVSPDGGQTIASGSDTKIKLWQTSDTGLTKITITGHEQTVNNVYFSPDGKN--L 663
 Qy 717 ATGSSDCFLKLWDLNQKCRNTMFGHTNSVNHCFSPDDKLLASCADGTLKLWDATSAN 776
 Db 664 ASASDHSIKLWDTTSGQLLMTLTGHSAGVITVRFSPDGTIAAGSEDKTVLMHRQDGK 723
 Qy 777 ERKSINVKQFFLNLEDPQEDMEVIVKCCSWSADGARIMVAANKIIFLFDIHTSGLLGEIH 836
 Db 724 LLKTLNGHQDWV-----SLSFSPDGKTLASASADKTIKLWRIADGKLVKTL 770
 Qy 837 TGHSTIOYCFSPQNHIAVALSOYCVELWNTDSRSKVADCRGHLWSWHGVFMFSPDGSS 896
 Db 771 KGHDSVWVDFNSDGAIAASARDNTIKLWNRHG-IELETPTHSGGGVYAVNFLPDSNI 829
 Qy 897 FLTSSDDOTIRLWE-----TKVKCNANVMLKQEVDFVQENEVWVLAVDHIRRELQIN 950
 Db 830 IASASLONTIRLWQRLPLSPILEVLNAGNSGVA-----VSFLHDGSIITAGADGNIQLWH 884
 Qy 951 GRTQIDYLTEAQVSCCCLS--PHLQYIAFGDENGAIIELELVNNRIIFQSRFOHKKTVWH 1008
 Db 885 SODGSLKTLPGNKAIYGISFTPGDGLIASANADKTVKIWRVROGKALKTLIGHDNEVVK 944
 Qy 1009 IOFTADEKTLISSDDASIQVNNWQDKCIF---LRGHOETVKDPRLLKNSRL--SWSFD 1064
 Db 945 VNFSPDGKTLASARDNTVKLWNVSDGK--FKKTLKGHTDEVFWVSFSPDGKIIASASAD 1002
 Qy 1065 GTVKVWNIITGNKEKDFVCHQGTVLSCDISHDATKFSSTADKTIKW-SFDLLPLPHL 1123
 Db 1003 KTIRLWDSFSGNLIKSLPAHNDLVYSVNFNPDGSMLASTADKTVKLWRSHDGL-LHTF 1061

Qy 1124 RHNCGVRCSAFSDVSTLLATGDNGEIRIWNVNGELLHLHCAPLSEBGAATHGGWVTDL 1183
 Db 1062 SGHSNVVYSSSPSPDGRYIASASEDKTVKIWDIG---HLLTTLTPQ-----HQAGVMSA 1112
 Qy 1184 CFSPDGKMLIS 1194
 Db 1113 IFSPDGKTLIS 1123

Search completed: January 27, 2005, 18:53:21
 Job time : 178 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 18:32:45 ; Search time 27 Seconds
(without alignments)
3065.365 Million cell updates/sec

Title: US-10-646-396-2
Perfect score: 6619
Sequence: 1 MDKARNCLLQHRLEAKDI.....FKTVTVVDNLGILYILQTL 1248

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6352.5	96.0	1205	3	US-09-092-508-16
2	6352.5	96.0	1205	3	US-09-435-115-16
3	6282	94.9	1194	3	US-09-092-508-2
4	6282	94.9	1194	3	US-09-435-115-2
5	6282	94.9	1194	3	US-09-069-023-26
6	6282	94.9	1194	4	US-09-098-310-2
7	6282	94.9	1194	4	US-09-538-092-825
8	408	6.2	2627	2	US-08-751-189-3
9	408	6.2	2627	2	US-09-060-836-3
10	408	6.2	2627	3	US-09-184-445-3
11	399.5	6.0	2629	2	US-08-751-189-4
12	399.5	6.0	2629	2	US-09-060-836-4
13	399.5	6.0	2629	3	US-09-184-445-4
14	322	4.9	251	4	US-09-291-170A-13
15	322	4.9	251	4	US-09-724-884-13
16	322	4.9	251	4	US-09-724-592-13
17	316	4.8	514	1	US-08-190-802A-66
18	316	4.8	514	3	US-08-477-346-66
19	316	4.8	514	3	US-08-473-089-66
20	316	4.8	514	4	US-08-487-072A-66
21	312	4.7	409	1	US-08-190-802A-51
22	312	4.7	409	3	US-08-477-346-51
23	312	4.7	409	3	US-08-473-089-51
24	312	4.7	409	4	US-08-487-072A-51
25	311.5	4.7	409	2	US-08-283-917-3
26	311.5	4.7	409	2	US-08-961-716-3
27	311.5	4.7	409	4	US-09-538-092-1119

28	311.5	4.7	410	2	US-08-283-917-9	Sequence 9, Appli
29	311.5	4.7	410	2	US-08-961-716-9	Sequence 9, Appli
30	309	4.7	318	1	US-08-190-802A-33	Sequence 33, Appli
31	309	4.7	318	3	US-08-477-346-33	Sequence 33, Appli
32	309	4.7	318	3	US-08-473-089-33	Sequence 33, Appli
33	309	4.7	318	4	US-08-487-072A-33	Sequence 33, Appli
34	307.5	4.6	316	4	US-09-828-310-12	Sequence 12, Appli
35	297.5	4.5	640	4	US-09-177-165A-30	Sequence 30, Appli
36	294	4.4	540	4	US-09-213-888-7	Sequence 7, Appli
37	294	4.4	540	4	US-09-213-888-10	Sequence 10, Appli
38	294	4.4	540	4	US-09-328-877D-7	Sequence 7, Appli
39	294	4.4	540	4	US-09-328-877D-10	Sequence 10, Appli
40	294	4.4	545	4	US-09-213-888-6	Sequence 6, Appli
41	294	4.4	545	4	US-09-328-877D-6	Sequence 6, Appli
42	294	4.4	553	4	US-09-213-888-5	Sequence 5, Appli
43	294	4.4	553	4	US-09-328-877D-5	Sequence 5, Appli
44	294	4.4	559	4	US-09-213-888-9	Sequence 9, Appli
45	294	4.4	559	4	US-09-328-877D-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-092-508-16
; Sequence 16, Application US/09092508
; Patent No. 6291643
; GENERAL INFORMATION:
; APPLICANT: Hensel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6291643west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/092,508
; FILING DATE: 05-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,807
; FILING DATE: 05-JUN-1997
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-09-092-508-16

Query Match 96.0%; Score 6352.5; DB 3; Length 1205;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1205; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

```

QY 1 MDKARNCLLOHREALEKDIKTSYIMDMHISDGFLTI SEEEKVNEPTQOORAAMLIKMI 60
DB 1 MDKARNCLLOHREALEKDIKTSYIMDMHISDGFLTI SEEEKVNEPTQOORAAMLIKMI 60
QY 61 LKKNDSVSVFYNALLHEGYKDLAALLHDCGIPWSSSSGKDSVSGITSYVVRTVLCGGVP 120
DB 61 LKKNDSVSVFYNALLHEGYKDLAALLHDCGIPWSSSSGKDSVSGITSYVVRTVLCGGVP 120
QY 121 QRPVVVTRKKLVNAIOOKLSKLKGPWVTHGHMAGCKSVLAABAVRSHLSLECCFPG 180
DB 121 QRPVVVTRKKLVNAIOOKLSKLKGPWVTHGHMAGCKSVLAABAVRSHLSLECCFPG 180
QY 181 GYHWSVQKDSGLIMKLNCLTRLDODSEFSORPLNIEENKORILMLKPHRSLL 240
DB 181 GYHWSVQKDSGLIMKLNCLTRLDODSEFSORPLNIEENKORILMLKPHRSLL 240
QY 241 ILDDVMSWVILKAFDSOQILLTRDKSVTDSVMGPKVVPVSSSLGKEGLBILSLFVN 300
DB 241 ILDDVMSWVILKAFDSOQILLTRDKSVTDSVMGPKVVPVSSSLGKEGLBILSLFVN 300
QY 301 MKKADLPQAHSHIIECKGSPVLSLIGALLRDPFRNWEYILKOLNQKQFRIRKSSSYD 360
DB 301 MKKADLPQAHSHIIECKGSPVLSLIGALLRDPFRNWEYILKOLNQKQFRIRKSSSYD 360
QY 361 YEALDEAMSISEVLRDIDKYVTDLSILQKQVPTKVLGILWDMETEEVEDILOEFVN 420
DB 361 YEALDEAMSISEVLRDIDKYVTDLSILQKQVPTKVLGILWDMETEEVEDILOEFVN 420
QY 421 KSLLCFDRNGKSFYRLHDLQVDFLETKNCSOLQDLHKKIITQFORVHQPHTLSPQOEDC 480
DB 421 KSLLCFDRNGKSFYRLHDLQVDFLETKNCSOLQDLHKKIITQFORVHQPHTLSPQOEDC 480
QY 481 MYWYFLAYHWSAKMKECALMFSIDWKATKELVGPALHIEFVEYRHLDEKCAV 540
DB 481 MYWYFLAYHWSAKMKECALMFSIDWKATKELVGPALHIEFVEYRHLDEKCAV 540
QY 541 SENFOEFLSINGHILGROPENIYVOLGCEPSETSVYQOAKLOKQVDMGMYLWINK 600
DB 541 SENFOEFLSINGHILGROPENIYVOLGCEPSETSVYQOAKLOKQVDMGMYLWINK 600
QY 601 KNTNLSRLVPRHTDAVTHACSEDOQRIASCGADKTLQVFKARTGEKLEIKAHEDV 660
DB 601 KNTNLSRLVPRHTDAVTHACSEDOQRIASCGADKTLQVFKARTGEKLEIKAHEDV 660
QY 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLATGS 720
DB 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLATGS 720
QY 721 SDCFLKLDLQKCECRNTMFGHTNSVNHCRFSPDDKLLASCSDGTLKLDATSANERKS 780
DB 721 SDCFLKLDLQKCECRNTMFGHTNSVNHCRFSPDDKLLASCSDGTLKLDATSANERKS 780
QY 781 INVKOFFNLNLEDPQEDMEVIVKCCSWASDAGRIWAAKNKIFLFDIHTSGILGHIHTGHH 840
DB 781 INVKOFFNLNLEDPQEDMEVIVKCCSWASDAGRIWAAKNKIFLFDIHTSGILGHIHTGHH 840
QY 841 STIQYCDSPQNLAVVALSYCVELWNTDSRSKVADCRGHLSSVHGVMPSPDGSFLTS 900
DB 841 STIQYCDSPQNLAVVALSYCVELWNTDSRSKVADCRGHLSSVHGVMPSPDGSFLTS 900
QY 901 SDOQIRLWETKVKCSAVALKQEDVUVFQENWVLAVIDHRLQILNCRGQDYLIT 960
DB 901 SDOQIRLWETKVKCSAVALKQEDVUVFQENWVLAVIDHRLQILNCRGQDYLIT 960
QY 961 EQAVSCCLSLPHQYIATGDBNGAIETLELVNRIPOSFOHKKTVWHIOPTADEKTLIS 1020
DB 961 EQAVSCCLSLPHQYIATGDBNGAIETLELVNRIPOSFOHKKTVWHIOPTADEKTLIS 1020
QY 1021 SSDDAEIQVNWQDLKCIPLRQHGTVDKDFRLLKNSRLLSWFGDTVKVNNIITGNKEKD 1080
DB 1021 SSDDAEIQVNWQDLKCIPLRQHGTVDKDFRLLKNSRLLSWFGDTVKVNNIITGNKEKD 1080
QY 1037 SSDDAEIQVNWQDLKCIPLRQHGTVDKDFRLLKNSRLLSWFGDTVKVNNIITGNKEKD 1037
DB 1037 SSDDAEIQVNWQDLKCIPLRQHGTVDKDFRLLKNSRLLSWFGDTVKVNNIITGNKEKD 1037

RESULT 2
US-09-435-115-16
; Sequence 16, Application US/09435115
; Patent No. 6346807
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAP-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 634607 West Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/435,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/092,508
; FILING DATE:
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-09-435-115-16

Query Match 96.0%; Score 6352.5; DB 3; Length 1205;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1205; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

QY 1 MDKARNCLLOHREALEKDIKTSYIMDMHISDGFLTI SEEEKVNEPTQOORAAMLIKMI 60
DB 1 MDKARNCLLOHREALEKDIKTSYIMDMHISDGFLTI SEEEKVNEPTQOORAAMLIKMI 60
QY 61 LKKNDSVSVFYNALLHEGYKDLAALLHDCGIPWSSSSGKDSVSGITSYVVRTVLCGGVP 120
DB 61 LKKNDSVSVFYNALLHEGYKDLAALLHDCGIPWSSSSGKDSVSGITSYVVRTVLCGGVP 120
QY 121 QRPVVVTRKKLVNAIOOKLSKLKGPWVTHGHMAGCKSVLAABAVRSHLSLECCFPG 180
DB 121 QRPVVVTRKKLVNAIOOKLSKLKGPWVTHGHMAGCKSVLAABAVRSHLSLECCFPG 180

```

```

Db      121  QRPVVVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCKSVLAABAVRDSHLLGCGPFG 180
Qy      181  GVHWSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEBAKDRRLILMLRKHPRSL 240
Db      181  GVHWSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEBAKDRRLILMLRKHPRSL 240
Qy      241  ILDDVWDSWLVKAFDSQOQILLTTTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFVN 300
Db      241  ILDDVWDSWLVKAFDSQOQILLTTTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFVN 300
Qy      301  MKKADLPQOASIIKECKGSPLVVSLIGALLRDFPNRWEYILKOLQKQFKRIRKSSVD 360
Db      301  MKKADLPQOASIIKECKGSPLVVSLIGALLRDFPNRWEYILKOLQKQFKRIRKSSVD 360
Qy      361  YEALDEAMSISVEMLEREDIKDYTTDLSILQKDVKVPVKVLCILMDMETEEVEDIILQEFVN 420
Db      361  YEALDEAMSISVEMLEREDIKDYTTDLSILQKDVKVPVKVLCILMDMETEEVEDIILQEFVN 420
Qy      421  KSLFLCDRNGKSFRIYLDLQVDFUTERNCSQLODLHKIITQFORHYQPHTLSPDQSDC 480
Db      421  KSLFLCDRNGKSFRIYLDLQVDFUTERNCSQLODLHKIITQFORHYQPHTLSPDQSDC 480
Qy      481  MYWYNFLAYHWASAKWHKELCALMFLDWIKAKTELVGPAHLIHEFVEYRHLIDBKCAV 540
Db      481  MYWYNFLAYHWASAKWHKELCALMFLDWIKAKTELVGPAHLIHEFVEYRHLIDBKCAV 540
Qy      541  SENFOEFLSLNGHLGRPFNIVOLGCEPETSEVYQOAKLQAKQEVNDGMLYLEWINK 600
Db      541  SENFOEFLSLNGHLGRPFNIVOLGCEPETSEVYQOAKLQAKQEVNDGMLYLEWINK 600
Qy      601  KNITNLSRLVVRPHPTDAVYHACFSDGQRIASCADKTLQVFKABTGKLEIKAHEDV 660
Db      601  KNITNLSRLVVRPHPTDAVYHACFSDGQRIASCADKTLQVFKABTGKLEIKAHEDV 660
Qy      661  LCCAFSTDTRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGS 720
Db      661  LCCAFSTDTRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLLATGS 720
Qy      721  SDCFLKLDLWLNQECRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 780
Db      721  SDCFLKLDLWLNQECRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 780
Qy      781  INVQOFFFLNLEDPQEDMEVIVKCCSWSDAGARIMVAANKIPLFDIHTSGLLGEIHTGHH 840
Db      781  INVQOFFFLNLEDPQEDMEVIVKCCSWSDAGARIMVAANKIPLFDIHTSGLLGEIHTGHH 840
Qy      841  STIQCYDFSPQNHAVLVALSQYCVLELWNTDSRSKVADCRGHLVWGHVYMFSPDGSSFLTS 900
Db      841  STIQCYDFSPQNHAVLVALSQYCVLELWNTDSRSKVADCRGHLVWGHVYMFSPDGSSFLTS 900
Qy      901  SDDQTLRLWETKVKCKNSAVMLKQEVVDVFOENVMVLAVDHIRLQLINGRTGQIDYLT 960
Db      901  SDDQTLRLWETKVKCKNSAVMLKQEVVDVFOENVMVLAVDHIRLQLINGRTGQIDYLT 960
Qy      961  EAQVSCCLSPHLOVIAFGDENGAEIILELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 1020
Db      961  EAQVSCCLSPHLOVIAFGDENGAEIILELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 1020
Qy      1021  SSDDAEIOVWNLQDKCIFLGHQHTVKDFRLLKNSRLLSWSFDGTVKVMNIIITGNKEKD 1080
Db      1021  SSDDAEIOVWNLQDKCIFLGHQHTVKDFRLLKNSRLLSWSFDGTVKVMNIIITGNKEKD 1080
Qy      1081  FVCHGTQVLSCDISHDATKFSSTADKTAKTWSFDLLLPLHELGRHNCVCSAFSVDST 1140
Db      1081  FVCHGTQVLSCDISHDATKFSSTADKTAKTWSFDLLLPLHELGRHNCVCSAFSVDST 1140
Qy      1141  LLATGDDNGEIRIWNVSGELLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1200
Db      1141  LLATGDDNGEIRIWNVSGELLHLCAPLSEGAATHGGWVTDLCFSPDGKMLISAGGYIK 1200
Qy      1201  WNWVVTGSSQTFYTNCTNKKIHSVDPFKTYVTVDNLGILYILOTLE 1248

```

```

Db      1158  WNWVVTGSSQTFYTNCTNKKIHSVDPFKTYVTVDNLGILYILOTLE 1205

RESULT 3
US-09-092-508-2
; Sequence 2, Application US/09092508
; Patent No. 6291643
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6291643west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/092,508
; FILING DATE: 05-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,807
; FILING DATE: 05-JUN-1997
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-09-092-508-2

```

```

Query Match      94.9%; Score 6282; DB 3; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

Qy      1  MDAKARNCLLOHRALEKDIKTSYIMDMHISDGFITISEEEKVNRNEPTQOORAAAMLIKMI 60
Db      1  MDAKARNCLLOHRALEKDIKTSYIMDMHISDGFITISEEEKVNRNEPTQOORAAAMLIKMI 60
Qy      61  LKKNDSDSVSYFNALLHEGYKDLAALLHDGIPVYSSSSGKDSVSGITSVYRTVLCGGVP 120
Db      61  LKKNDSDSVSYFNALLHEGYKDLAALLHDGIPVYSSSSGKDSVSGITSVYRTVLCGGVP 109
Qy      121  QRPVVVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCKSVLAABAVRDSHLLGCGPFG 180
Db      110  QRPVVVTRKKLVNAIQOKLSKLGEPGWVTHGMAGCKSVLAABAVRDSHLLGCGPFG 169
Qy      181  GVHWSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEBAKDRRLILMLRKHPRSL 240
Db      170  GVHWSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEBAKDRRLILMLRKHPRSL 229
Qy      241  ILDDVWDSWLVKAFDSQOQILLTTTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFVN 300
Db      230  ILDDVWDSWLVKAFDSQOQILLTTTRDKSVTDSVMGPKVVPVSSSLGKEGLEILSLFVN 289

```

```

QY 301 MKKADLPEQAHSIIKECKGSPVLSLIGALLRFPNRMWYLLKQLONQKQKRIKSSSYD 360
DB 290 MKKADLPEQAHSIIKECKGSPVLSLIGALLRFPNRMWYLLKQLONQKQKRIKSSSYD 349
QY 361 YEALDEAMSISVEMLRREDIKDYITDLSILQKQKVPKVLKILMDMETEVEDILQEFVN 420
DB 350 YEALDEAMSISVEMLRREDIKDYITDLSILQKQKVPKVLKILMDMETEVEDILQEFVN 409
QY 421 KSLFLCDRNGKSFYRLHDLQVDFTEKNGSQDLHKKIITQFORHYHQPHILSPQEDC 480
DB 410 KSLFLCDRNGKSFYRLHDLQVDFTEKNGSQDLHKKIITQFORHYHQPHILSPQEDC 469
QY 481 MYWYNFLAYHMASAKWHELCALMFSIDWKATKELVGPALHIFEFVEYHILDEKCAV 540
DB 470 MYWYNFLAYHMASAKWHELCALMFSIDWKATKELVGPALHIFEFVEYHILDEKCAV 529
QY 541 SENFQFLSLNGHLGROPPFNIVQLGCEPTESEVQOAKLOAQEVDNGMLYLEWINK 600
DB 530 SENFQFLSLNGHLGROPPFNIVQLGCEPTESEVQOAKLOAQEVDNGMLYLEWINK 589
QY 601 KNTNLRLVVRPHDTAVYHACFSEDEQRIASCADKTLQVPKATGKLELKAHEDEV 660
DB 590 KNTNLRLVVRPHDTAVYHACFSEDEQRIASCADKTLQVPKATGKLELKAHEDEV 649
QY 661 LCCAFSTDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLLATGS 720
DB 650 LCCAFSTDRFATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHLLLATGS 709
QY 721 SDCFLKLDLNOKECRNMTFNTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
DB 710 SDCFLKLDLNOKECRNMTFNTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 769
QY 781 INVQKQFLNLEDPQEDMEVIKCCSWSADGARIMVAANKIPLFDIHTSGLLGEIHTGHH 840
DB 770 INVQKQFLNLEDPQEDMEVIKCCSWSADGARIMVAANKIPLFDIHTSGLLGEIHTGHH 811
QY 841 STIOYCDPSPQNHAAVVALSQYVELWNTDSRSKVADCRGHLSSWVHGVMFSPDGSFLTS 900
DB 812 STIOYCDPSPQNHAAVVALSQYVELWNTDSRSKVADCRGHLSSWVHGVMFSPDGSFLTS 846
QY 901 SDDOTIRLWETKVCNSAVMLKQEDVVFQENWVLAVIDHRLQOLINGRTGDDYLT 960
DB 847 SDDOTIRLWETKVCNSAVMLKQEDVVFQENWVLAVIDHRLQOLINGRTGDDYLT 906
QY 961 EAQVSCCLSPHLQYIARFGDENGAEIIELVNRRIPQSRFQHKTVWHIQTADKTLIS 1020
DB 907 EAQVSCCLSPHLQYIARFGDENGAEIIELVNRRIPQSRFQHKTVWHIQTADKTLIS 966
QY 1021 SSDAIEIQWNNQDLKCIFLGHQETVQDFRLLKNSRLLSWSPDGTWKVWNIIITGNKEKD 1080
DB 967 SSDAIEIQWNNQDLKCIFLGHQETVQDFRLLKNSRLLSWSPDGTWKVWNIIITGNKEKD 1026
QY 1081 FVCHQGTVLSCDISHDATKFSSTASDKTAKIWSFDLPLHLRGNHGCVRCSAFSDST 1140
DB 1027 FVCHQGTVLSCDISHDATKFSSTASDKTAKIWSFDLPLHLRGNHGCVRCSAFSDST 1086
QY 1141 LLATGDNGEIRIWNVNGELLHLCAPLSEGAATHGGVWTDLCFSPDGKMLISAGYIK 1200
DB 1087 LLATGDNGEIRIWNVNGELLHLCAPLSEGAATHGGVWTDLCFSPDGKMLISAGYIK 1146
QY 1201 WNNVVTGESSOTFYNGNGLKKHVSPDKTYVVDNLGILYILQTL 1248
DB 1147 WNNVVTGESSOTFYNGNGLKKHVSPDKTYVVDNLGILYILQTL 1194

```

RESULT 4

US-09-435-115-2

; Sequence 2, Application US/09435115

; Patent No. 6346607

; GENERAL INFORMATION:

; APPLICANT: Henzel, William J.

; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3

```

; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6346607 West Center, 90 South Seventh St.
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-Seq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/435,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/092,508
; FILING DATE:
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-09-435-115-2

```

Query Match 94.9%; Score 6282; DB 3; Length 1194;

Best Local Similarity 95.7%; Pred. No. 0;

Matches 194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

```

QY 1 MDKARNCLLOHREALEKDKITSYIMDMISDGLTISEEEKVNRNEPTQOORAMLIKMI 60
DB 1 MDKARNCLLOHREALEKDKITSYIMDMISDGLTISEEEKVNRNEPTQOORAMLIKMI 60
QY 61 LKQNDSTSVFYNALHEGYKDLAALLHDGIPVSSSSGKDSVSGITSYVTVLCEGVP 120
DB 61 LKQNDSTSVFYNALHEGYKDLAALLHDGIPVSSSSGKDSVSGITSYVTVLCEGVP 109
QY 121 QRPVVFTRKLVNAIQCKLSKLKGEPEGWVTIHGMAGCGKSVLAABAVRDHSLLEGCPG 180
DB 110 QRPVVFTRKLVNAIQCKLSKLKGEPEGWVTIHGMAGCGKSVLAABAVRDHSLLEGCPG 169
QY 181 GVHWVSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEEAKDRLRILMLRKHPRSL 240
DB 170 GVHWVSVGKQKSGLLMKLQNLCTRLDQDESFSQRLPLNIEEAKDRLRILMLRKHPRSL 229
QY 241 ILDDVDSWVLKAFDSCQCIILTTTRDKSVTDVSMGPKVYVVPVSSSLCKEKGLEILSLFVN 300
DB 230 ILDDVDSWVLKAFDSCQCIILTTTRDKSVTDVSMGPKVYVVPVSSSLCKEKGLEILSLFVN 289
QY 301 MKKADLPEQAHSIIKECKGSPVLSLIGALLRDPNRMWYLLKQLONQKQKRIKSSSYD 360
DB 290 MKKADLPEQAHSIIKECKGSPVLSLIGALLRDPNRMWYLLKQLONQKQKRIKSSSYD 349
QY 361 YEALDEAMSISVEMLRREDIKDYITDLSILQKQKVPKVLKILMDMETEVEDILQEFVN 420
DB 350 YEALDEAMSISVEMLRREDIKDYITDLSILQKQKVPKVLKILMDMETEVEDILQEFVN 409
QY 421 KSLFLCDRNGKSFYRLHDLQVDFTEKNGSQDLHKKIITQFORHYHQPHILSPQEDC 480

```

Db 410 KSLFLCDNRGKSFYIYLDQVDFLTERKNSQLDLHKKIITQFORHQPHLSPQDSDC 469
Qy 481 MYWNYFLAYHMASAKMHELCALMFSLDWIKAKTELAVGPAHLIHFVEYRHLIDKDCAV 540
Db 470 MYWNYFLAYHMASAKMHELCALMFSLDWIKAKTELAVGPAHLIHFVEYRHLIDKDCAV 529
Qy 541 SENFOEFLSLNGHLLGRQPPFNIVOLGLCEPSETSEYVQAKLQAKQEVNDGMLYLEWINK 600
Db 530 SENFOEFLSLNGHLLGRQPPFNIVOLGLCEPSETSEYVQAKLQAKQEVNDGMLYLEWINK 589
Qy 601 KNIITNLSRLVVRPHTDAVTHACFSEDGQRIASCADKTLQVFKAETGKLEIKAHEDV 660
Db 590 KNIITNLSRLVVRPHTDAVTHACFSEDGQRIASCADKTLQVFKAETGKLEIKAHEDV 649
Qy 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS 720
Db 650 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS 709
Qy 721 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
Db 710 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 769
Qy 781 INVKQFFLNLEDPQEDMEVIVKCCSWASADGARIIMVAANKIFLFDIHTSGLLGEIHTGHH 840
Db 770 INVKQFFLNLEDPQEDMEVIVKCCSWASADGARIIMVAANKIF- 811
Qy 841 STIOYCFSPQNHVALVALSOYVELWNTDSRSKVADCRGHLWSVGHVGMFSPDGSFPTS 900
Db 812 -----LWNTDSRSKVADCRGHLWSVGHVGMFSPDGSFPTS 846
Qy 901 SDDQITRLWETKVKCKNSAVMLKQEVNVFOENVMVLAVDHIRLQLINGRTGQIDYLT 960
Db 847 SDDQITRLWETKVKCKNSAVMLKQEVNVFOENVMVLAVDHIRLQLINGRTGQIDYLT 906
Qy 961 EAOVSCCLSPHLQYIAFGDENGATEILELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 1020
Db 907 EAOVSCCLSPHLQYIAFGDENGATEILELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 966
Qy 1021 SSDDAEIQVWNNQDLKCIPLRGHQBTVKDFLLKNSRLLSWDFCTVKNWNIITGNKEK 1080
Db 967 SSDDAEIQVWNNQDLKCIPLRGHQBTVKDFLLKNSRLLSWDFCTVKNWNIITGNKEK 1026
Qy 1081 FVCHQGTVLSCDISHDATKFSSTSDAKTIWSFDLLPLHLRGNHGCVRCSAFSDST 1140
Db 1027 FVCHQGTVLSCDISHDATKFSSTSDAKTIWSFDLLPLHLRGNHGCVRCSAFSDST 1086
Qy 1141 LLATGDDNGBEIRIWNVSGELLHLCAPLSEGAATHGGVWTDLCFSPDGKMLISAGGYIK 1200
Db 1087 LLATGDDNGBEIRIWNVSGELLHLCAPLSEGAATHGGVWTDLCFSPDGKMLISAGGYIK 1146
Qy 1201 WNNVVTGESSOTFYTNGTNLKKIHVSPDKTAVTVDNIGILYILOTLE 1248
Db 1147 WNNVVTGESSOTFYTNGTNLKKIHVSPDKTAVTVDNIGILYILOTLE 1194

RESULT 5
US-09-069-023-26
; Sequence 26, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 1194

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-26
Query Match 94.9%; Score 6282; DB 3; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;
Qy 1 MDKARNCLLQHREALEKDIKTSYIMDHISDGLFIISEEEKVRNEPTQQQRAAMLKMI 60
Db 1 MDKARNCLLQHREALEKDIKTSYIMDHISDGLFIISEEEKVRNEPTQQQRAAMLKMI 60
Qy 61 LKKNDSVTSFYNNALLHGGYKDLAALLHDGIPVSSSSSGKDSVSGITSVTVLCEGGVP 120
Db 61 LKKNDSVTSFYNNALLHGGYKDLAALLHDGIPVSSSSSGKDSVSGITSVTVLCEGGVP 109
Qy 121 QRPVVVTRKKLVNAIQOKLSKLKGEPCGWVTHGMAGCGKSVLAEEAVRDSHLLSGCPFG 180
Db 110 QRPVVVTRKKLVNAIQOKLSKLKGEPCGWVTHGMAGCGKSVLAEEAVRDSHLLSGCPFG 169
Qy 181 GVHVSVGKQDKSGLLMKQLNLCRLDQDESFSQRLPLNIEAKDRLILMLRKHPRLSLL 240
Db 170 GVHVSVGKQDKSGLLMKQLNLCRLDQDESFSQRLPLNIEAKDRLILMLRKHPRLSLL 229
Qy 241 ILDDVMSWLVKAFDSQCIILLTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN 300
Db 230 ILDDVMSWLVKAFDSQCIILLTRDKSVTDSVMGPKYVVPVSSSLGKEGLEILSLFVN 289
Qy 301 MKKADLPEQAHSIKECKGSPVLSLIGALLRDPNRMWEYVILKOLNKFQKRIKSSSYD 360
Db 290 MKKADLPEQAHSIKECKGSPVLSLIGALLRDPNRMWEYVILKOLNKFQKRIKSSSYD 349
Qy 361 YEALDEAMSIIVEMLRREDIKDYTDLSILOKDVKPTKVLGILMDMETEEVEDILOEFVN 420
Db 350 YEALDEAMSIIVEMLRREDIKDYTDLSILOKDVKPTKVLGILMDMETEEVEDILOEFVN 409
Qy 421 KSLFLCDNRGKSFYIYLDQVDFLTERKNSQLDLHKKIITQFORHQPHLSPQDSDC 480
Db 410 KSLFLCDNRGKSFYIYLDQVDFLTERKNSQLDLHKKIITQFORHQPHLSPQDSDC 469
Qy 481 MYWNYFLAYHMASAKMHELCALMFSLDWIKAKTELAVGPAHLIHFVEYRHLIDKDCAV 540
Db 470 MYWNYFLAYHMASAKMHELCALMFSLDWIKAKTELAVGPAHLIHFVEYRHLIDKDCAV 529
Qy 541 SENFOEFLSLNGHLLGRQPPFNIVOLGLCEPSETSEYVQAKLQAKQEVNDGMLYLEWINK 600
Db 530 SENFOEFLSLNGHLLGRQPPFNIVOLGLCEPSETSEYVQAKLQAKQEVNDGMLYLEWINK 589
Qy 601 KNIITNLSRLVVRPHTDAVTHACFSEDGQRIASCADKTLQVFKAETGKLEIKAHEDV 660
Db 590 KNIITNLSRLVVRPHTDAVTHACFSEDGQRIASCADKTLQVFKAETGKLEIKAHEDV 649
Qy 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS 720
Db 650 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGS 709
Qy 721 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 780
Db 710 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDATSANERKS 769
Qy 781 INVKQFFLNLEDPQEDMEVIVKCCSWASADGARIIMVAANKIFLFDIHTSGLLGEIHTGHH 840
Db 770 INVKQFFLNLEDPQEDMEVIVKCCSWASADGARIIMVAANKIF- 811
Qy 841 STIOYCFSPQNHVALVALSOYVELWNTDSRSKVADCRGHLWSVGHVGMFSPDGSFPTS 900
Db 812 -----LWNTDSRSKVADCRGHLWSVGHVGMFSPDGSFPTS 846
Qy 901 SDDQITRLWETKVKCKNSAVMLKQEVNVFOENVMVLAVDHIRLQLINGRTGQIDYLT 960
Db 847 SDDQITRLWETKVKCKNSAVMLKQEVNVFOENVMVLAVDHIRLQLINGRTGQIDYLT 906
Qy 961 EAOVSCCLSPHLQYIAFGDENGATEILELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 1020
Db 907 EAOVSCCLSPHLQYIAFGDENGATEILELVNRRIFQSRFQHKKTVMHIQFTADEKTLIS 966
Qy 1021 SSDDAEIQVWNNQDLKCIPLRGHQBTVKDFLLKNSRLLSWDFCTVKNWNIITGNKEK 1080
Db 967 SSDDAEIQVWNNQDLKCIPLRGHQBTVKDFLLKNSRLLSWDFCTVKNWNIITGNKEK 1026
Qy 1081 FVCHQGTVLSCDISHDATKFSSTSDAKTIWSFDLLPLHLRGNHGCVRCSAFSDST 1140
Db 1027 FVCHQGTVLSCDISHDATKFSSTSDAKTIWSFDLLPLHLRGNHGCVRCSAFSDST 1086
Qy 1141 LLATGDDNGBEIRIWNVSGELLHLCAPLSEGAATHGGVWTDLCFSPDGKMLISAGGYIK 1200
Db 1087 LLATGDDNGBEIRIWNVSGELLHLCAPLSEGAATHGGVWTDLCFSPDGKMLISAGGYIK 1146
Qy 1201 WNNVVTGESSOTFYTNGTNLKKIHVSPDKTAVTVDNIGILYILOTLE 1248
Db 1147 WNNVVTGESSOTFYTNGTNLKKIHVSPDKTAVTVDNIGILYILOTLE 1194


```

Db 907 EAQVSCCLSPHLOQYIAFGDNGAIELELVNRRIFQSRFOHKTVMHIQFTADEKTLIS 966
Qy 1021 SSDASIQVNNQWLDKICIFLRGHQETVKDPRLLKNSRLLSWSFGDTGVKVMNIIITGNKEKD 1080
Db 967 SSDASIQVNNQWLDKICIFLRGHQETVKDPRLLKNSRLLSWSFGDTGVKVMNIIITGNKEKD 1026
Qy 1081 FVCHQGTVLSCDISHDARTKFSSTADTKAKIWSFDLPLHLRGHNCVRCSAFSDST 1140
Db 1027 FVCHQGTVLSCDISHDARTKFSSTADTKAKIWSFDLPLHLRGHNCVRCSAFSDST 1086
Qy 1141 LLATGDDNGEIRIWNVSNGLLHLCAPLSEGAATHGGWTDLCFSPDGKMLISAGGYIK 1200
Db 1087 LLATGDDNGEIRIWNVSNGLLHLCAPLSEGAATHGGWTDLCFSPDGKMLISAGGYIK 1146
Qy 1201 WNVVTGESSQTFYNTGNLKKIHVSPDKTYVTVVNDLGLIYILOTLE 1248
Db 1147 WNVVTGESSQTFYNTGNLKKIHVSPDKTYVTVVNDLGLIYILOTLE 1194

RESULT 6
US-09-098-310-2
; Sequence 2, Application US/09098310
; Patent No. 6403765
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: TRUNCATED APAF-1 AND METHODS OF USE
; FILE REFERENCE: 480140.438
; CURRENT APPLICATION NUMBER: US/09/098,310
; PRIORITY FILING DATE: 1998-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1194
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-098-310-2

Query Match 94.9%; Score 6282; DB 4; Length 1194;
Best Local Similarity 95.7%; Pred No 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

Qy 1 MDKARNCLLOHREALEKDIKTSYIMDMISDGLTISEEKVNRNEPTQOORAMLIKMI 60
Db 1 MDKARNCLLOHREALEKDIKTSYIMDMISDGLTISEEKVNRNEPTQOORAMLIKMI 60
Qy 61 LKDNDSYVSYFNALLHGYKDLAALLHGDGIPVWSSSGKDSVSGITSYRTVLCRGGVP 120
Db 61 LKDNDSYVSYFNALLHGYKDLAALLHGDGIPVWSSSGKDSVSGITSYRTVLCRGGVP 109
Qy 121 QRPVVFVTRKKLVNAIQKLSKLKGEFGWVTHGMAGCKSVLAEEAVRDSHLLGCGFFG 180
Db 110 QRPVVFVTRKKLVNAIQKLSKLKGEFGWVTHGMAGCKSVLAEEAVRDSHLLGCGFFG 169
Qy 181 GVHWVSQKQSGILLMKLQNLCTRLDQDESFSQRLPLNIEBEAKDRILMLRKHPRSLL 240
Db 170 GVHWVSQKQSGILLMKLQNLCTRLDQDESFSQRLPLNIEBEAKDRILMLRKHPRSLL 229
Qy 241 ILDDVWDSWLVKAFDSQOILLTRDKSVTDSVMGPKVVPVSSSLGKGLSILFVN 300
Db 230 ILDDVWDSWLVKAFDSQOILLTRDKSVTDSVMGPKVVPVSSSLGKGLSILFVN 289
Qy 301 MKKADLPEQAHSIIECKGSPVLSLIGALLRDFNREYVYLKQNKQPKRIRKSSSD 360
Db 290 MKKADLPEQAHSIIECKGSPVLSLIGALLRDFNREYVYLKQNKQPKRIRKSSSD 349
Qy 361 YEALDEMSISVEMREDIKDYTDLSILQKQVPTKVLCLMDMTEVEEDILQEFVN 420
Db 350 YEALDEMSISVEMREDIKDYTDLSILQKQVPTKVLCLMDMTEVEEDILQEFVN 409
Qy 421 KSLILCDRNGKSFYRILHDLQVDFLTERKNSQLDLHKKIITQFQRYHQPHTLSPQEDC 480

```

```

Db 410 KSLILCDRNGKSFYRILHDLQVDFLTERKNSQLDLHKKIITQFQRYHQPHTLSPQEDC 469
Qy 481 MYWYNFLAYHMASAKMHELCALMFSLDWIKAKTELVGPAHLIHEFVEYRHILDEKDCAV 540
Db 470 MYWYNFLAYHMASAKMHELCALMFSLDWIKAKTELVGPAHLIHEFVEYRHILDEKDCAV 529
Qy 541 SENFOEFLSINGHLGRQPPNIVOLGCEPETSSEVYQOAKLOAKQVNDGMLYLEWINK 600
Db 530 SENFOEFLSINGHLGRQPPNIVOLGCEPETSSEVYQOAKLOAKQVNDGMLYLEWINK 589
Qy 601 KNITNLSELVVRPHTDAVHACSESDGQRIASCAGADTKLOVFKAETGKLEIKAHEDV 660
Db 590 KNITNLSELVVRPHTDAVHACSESDGQRIASCAGADTKLOVFKAETGKLEIKAHEDV 649
Qy 661 LCCARSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLATGS 720
Db 650 LCCARSTDDRIATCSVDKVKIWNMTGELVHTYDEHSEQVNCCHFTNSSHLLATGS 709
Qy 721 SDCFLKLDLNOKECRNMTFPGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 780
Db 710 SDCFLKLDLNOKECRNMTFPGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 769
Qy 781 INVKOFFFLNLEDPQEDMEVIVKCCWSADGARIIMVAAKNKIFLFDIHTSGLLGEIHTGHH 840
Db 770 INVKOFFFLNLEDPQEDMEVIVKCCWSADGARIIMVAAKNKIFLFDIHTSGLLGEIHTGHH 811
Qy 841 STIOYCDSPQNHLAVVALSYQVCELVNMTDSRSKVDACRGLHSVWVGWVMSPDGSSFLTS 900
Db 812 STIOYCDSPQNHLAVVALSYQVCELVNMTDSRSKVDACRGLHSVWVGWVMSPDGSSFLTS 846
Qy 901 SDDQIRLWETKVKCKNSAVMLKQEVVDVVFQENEVWVLAVDHRRQLQINGRTGQIDYLT 960
Db 847 SDDQIRLWETKVKCKNSAVMLKQEVVDVVFQENEVWVLAVDHRRQLQINGRTGQIDYLT 906
Qy 961 EAQVSCCLSPHLOQYIAFGDNGAIELELVNRRIFQSRFOHKTVMHIQFTADEKTLIS 1020
Db 907 EAQVSCCLSPHLOQYIAFGDNGAIELELVNRRIFQSRFOHKTVMHIQFTADEKTLIS 966
Qy 1021 SSDASIQVNNQWLDKICIFLRGHQETVKDPRLLKNSRLLSWSFGDTGVKVMNIIITGNKEKD 1080
Db 967 SSDASIQVNNQWLDKICIFLRGHQETVKDPRLLKNSRLLSWSFGDTGVKVMNIIITGNKEKD 1026
Qy 1081 FVCHQGTVLSCDISHDARTKFSSTADTKAKIWSFDLPLHLRGHNCVRCSAFSDST 1140
Db 1027 FVCHQGTVLSCDISHDARTKFSSTADTKAKIWSFDLPLHLRGHNCVRCSAFSDST 1086
Qy 1141 LLATGDDNGEIRIWNVSNGLLHLCAPLSEGAATHGGWTDLCFSPDGKMLISAGGYIK 1200
Db 1087 LLATGDDNGEIRIWNVSNGLLHLCAPLSEGAATHGGWTDLCFSPDGKMLISAGGYIK 1146
Qy 1201 WNVVTGESSQTFYNTGNLKKIHVSPDKTYVTVVNDLGLIYILOTLE 1248
Db 1147 WNVVTGESSQTFYNTGNLKKIHVSPDKTYVTVVNDLGLIYILOTLE 1194

RESULT 7
US-09-538-092-825
; Sequence 825, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIORITY FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqformatter Version 0.9
; SEQ ID NO 825

```

```

; LENGTH: 1194
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)-(0)
; OTHER INFORMATION: Polypeptide Accession Number O14727
US-09-538-092-825

Query Match          94.9%; Score 6282; DB 4; Length 1194;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1194; Conservative 0; Mismatches 0; Indels 54; Gaps 2;

Qy 1 MDAKARNCLLQHREALEKDIKTSYIMDMHISDGLFTTISEEKVKRNEPTQQQRAAMLIKMI 60
Db 1 MDAKARNCLLQHREALEKDIKTSYIMDMHISDGLFTTISEEKVKRNEPTQQQRAAMLIKMI 60

Qy 61 LKKNDSSVSVFNALLHEGYKDLAALHDGIPVVSSSGKDSVSGITSYVTVLCEGVP 120
Db 61 LKKNDSSVSVFNALLHEGYKDLAALHDGIPVVSSSS-----VRTVLCBGGVP 109

Qy 121 QRPVVFTVRKKLVNAIQOKLSKLKGEPCGWVTIHGMAGCGKSVLAEEAVRDHSLLEGCPFG 180
Db 110 QRPVVFTVRKKLVNAIQOKLSKLKGEPCGWVTIHGMAGCGKSVLAEEAVRDHSLLEGCPFG 169

Qy 181 GVHVSVGQKDGSLMLKQLNCTRLDQESFSQRLPLNIEEAKDRILMLRKHPRSLL 240
Db 170 GVHVSVGQKDGSLMLKQLNCTRLDQESFSQRLPLNIEEAKDRILMLRKHPRSLL 229

Qy 241 ILDDVMSWLVKAPDSQOQIILTRDKSVTDSVNGPKVVPVSSSLGKEGLEILSLFVN 300
Db 230 ILDDVMSWLVKAPDSQOQIILTRDKSVTDSVNGPKVVPVSSSLGKEGLEILSLFVN 289

Qy 301 MKKADLPQAHSIIEKCGSLVLSLIGALLRDPNREYVYLKOLQKQFRIKSSSYD 360
Db 290 MKKADLPQAHSIIEKCGSLVLSLIGALLRDPNREYVYLKOLQKQFRIKSSSYD 349

Qy 361 YEALDEAMSIIVEMLRDIDKYTDLSTLQKDVKVPYTKVLCILMDMETEVEYDLOEFVN 420
Db 350 YEALDEAMSIIVEMLRDIDKYTDLSTLQKDVKVPYTKVLCILMDMETEVEYDLOEFVN 409

Qy 421 KSLFLCDRNGKSPRYLHDLQVDFLTKNCSQLODLHKIITQORHQPHTLSPDQEDC 480
Db 410 KSLFLCDRNGKSPRYLHDLQVDFLTKNCSQLODLHKIITQORHQPHTLSPDQEDC 469

Qy 481 MYTNFLAYHMASAKMHELCALMFLSDWIKAETELVGPALHIEFVEYRHLIDKCAV 540
Db 470 MYTNFLAYHMASAKMHELCALMFLSDWIKAETELVGPALHIEFVEYRHLIDKCAV 529

Qy 541 SENFQFLSLNGLLGRQFPFNIVOLGLCEPETSEVYQQAQKQAEVDNGLYLEWINK 600
Db 530 SENFQFLSLNGLLGRQFPFNIVOLGLCEPETSEVYQQAQKQAEVDNGLYLEWINK 589

Qy 601 KNIITNLSRLVVRPHTDAVYHACFSDGQRIASCADKTLQVFKAETGEKLEIKAHEDV 660
Db 590 KNIITNLSRLVVRPHTDAVYHACFSDGQRIASCADKTLQVFKAETGEKLEIKAHEDV 649

Qy 661 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYVDEHSEQVNCCHFTNSHLLLATGS 720
Db 650 LCCAFSTDDRFIATCSVDKVKIWNMTGELVHTYVDEHSEQVNCCHFTNSHLLLATGS 709

Qy 721 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 780
Db 710 SDCFLKLDLWLNQKCRNTMFGHTNSVNHCRFSPDDKLLASCADGTLKLDWATSANERKS 769

Qy 781 INVQKFFLNLEDPOEDMEVIVKCCSWSDAGARIMVAANKKIFLPDIHTSGLLGEIHTGHH 840
Db 770 INVQKFFLNLEDPOEDMEVIVKCCSWSDAGARIMVAANKKIF----- 811

Qy 841 STIQYCDFSPQNHILAVVALSYCVELWNTDTSRKSVDACRGLHSVHGVMFSPDGSSFLT 900
Db 812 -----LWNTDTSRKSVDACRGLHSVHGVMFSPDGSSFLT 846

```

RESULT 8

```

US-08-751-189-3
; Sequence 3, Application US/08751189
; Patent No. 5919656
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 5919656el Genes Encoding Telomerase Protein
; TITLE OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08751,189
; APPLICATION NUMBER: US/08751,189
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleksi, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2627 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-751-189-3

Query Match          6.2%; Score 408; DB 2; Length 2627;
Best Local Similarity 21.0%; Pred. No. 4.1e-27;
Matches 296; Conservative 170; Mismatches 492; Indels 454; Gaps 58;

Qy 97 SSGKDSVSGITSYVTVLCE-----GGVQRPVVVTRKLVNAIQOKLSK-- 142
Db 1083 AAGRPYVGGLEEFQQLVLDVWNNIQLYLOPQALLEGQ-VSIPDDDLVQATFQOLQKPP 1141

Qy 143 -----LKGEPGWVTTHGMAGCGKSVLAEEAVRDHSLLEGCPGGV--- 182

```


183 QY -----HWVSVGKQKSGLLMKLQNLCTRLDQDESFQRLPLNIEBAKDRILMLRKHPR 237
1195 DB XLVPFHP-SGARPQGLALTLLRLCTYLRGOLKEPFGALPSTYRSLVWELQORLLPKSAE 1253
238 QY -----LILD-----DWDWSWLKAFDSQOQILLTTTRDKSVTDSVMPKVVV 280
1254 DB SLHPQOTQVLIIDGADRLVDQNGQLISDWIPKLPKPCVHLVLSV-----SAGLG----- 1304
281 QY PVESLGKKGLEILSLFVNMMKADLPQAHSAIIECKGSPVLVSLIGALLRDPFNWREY 340
1305 DB -----ETLEQSQGAHVLA-----GPLEASARARLVRE-----ELAL 1336
341 QY YLKQLQNKQFPR-----IRKSSSYDYEALDEAMSIIVEMLREDIKYDTLDSILQKDVK 394
1337 DB YGKRLESPFNQOMRLLLVTKRESGRPLYRLVTDHLRLFTLYEQVSERLRTL-----PAT 1391
395 QY VPTKVLCTILDMETEVEDIILQEF-----VNKSLFLCDR----- 428
1392 DB VPLLLOHILSLTEKEHGPVLIPOALTALVTRSGLTVDQLHGVLVSVRTLPGTKYSWEEA 1451
429 QY -----NGKSPRY-----LHDLDQVDFLTKNGSQLODLHKIITQFYHOPHTLSP 475
1452 DB VAAGNSGDPYMPGPPACLVQSLRSILGRPLERPGARCLPDGPLRTAAKRCYGR---P 1508
476 QY DOEDCMY-----W-----YNFLAYHASAKWHKELCALMPSLDWI 510
1509 DB GLEDTAHTILIAAQLWKTCDADASGTFTRSCPPEALGDLFPYHLLQSGNRGLLSKFTLNLHV 1568
511 QY KAKTELUGPAHLIHEFVY-----RHILDEKCAVSENFQEPFLSLNGHLHGQPPNI 563
1569 DB AAHLELGLVSLLENAHLYASVPKEBQKLEADVAV---FRTFLRQOASILSQ---YPLR 1623
564 QY V-QLGLCBPETSSEVYQAKLOAK-----QEVNMGMLYLEWINK-----KNITNLS-RLVVRP 613
1624 DB LPQQAANQLDPSPLCHQASLSLRWHLOHT-----LAWLNKPRMTKMQQSSSLSLAVSS 1677
614 QY HTDVAHYACFSEDGRIASCADKTLQVFKAEFTGEKLEIKAHEDVLCACAFSTDDRIA 673
1678 DB SPTAV---AFSTNGORAAVGTANGTVLLDRTQOEKSVSVSGCDISACILFLSDDTLFL 1734
674 QY TCSVDKVKYIWNMTGELVHYDEHSEQVNCCHFTNSSHLLLATGSSDCFLKWLNDLQK 733
1735 DB T-----AFDGLLELWDL-QH 1748
734 QY ECRNTMF-GHTNSVNHCRFSPDKLLASCADGTLKLD-----ATSANERKSNVVKQF 786
1749 DB GCRVLQTKAHQYQITGCCLSPCDLLATVCLGGCLKLDVTRGQLAFQHTYTPKSLNCVAF 1808
787 QY FLNLEDPQEDMEVIVKCCWS-----ADGARIMVAANKIIFLPIHTSGLLGEIHTGH 839
1809 DB -----HPEGQ---VIATGSWAGSISFFQVDGLKVTK-----DLGAPG----- 1842
840 QY HSTIQYCDPSQNHILAVVALSQCYELWNTDSRKVADCRGHLVSVHGMFSPDGSSFLT 899
1843 DB ASIRTLAFNVPGGVAVGRLDMSVLMWREGARLAAPPAHGGFVAALFLHAGCQLLT 1901
900 QY SSSDQITRLWE-----TKKVCNKSAMLKQEVDDVVVFQENVMVLAVDHIRRLQ 948
1902 DB AGEDGKQVWGSGLRPRGHLSLSLSPALSVALSPDGRV-----AVGYRADGIRIYKI 1956
949 QY INGRGTQIDYTEAQVSCC-LSPHLQYIAFGDENGAIETLEL-----VNNRIFOSRFOHK 1003
1957 DB SSGSQGAQCALDVAVALAWLSP--KVLVSGAEDGSLQGWALKECSLQSLWLLSRFQ-- 2012
1004 QY KTVWHIQTADEKTLISSDDAEIQVNWOL-----DKC-IFLGHQETV----- 1047
2013 DB KPV--LGLATSOELASASEFTVLWPRQLLTRPHKAEDPPCGTELGHGPGVSCCSFS 2070
1048 QY -----KDFRLL-----KNSRLLSWSFDGTV 1067
2071 DB TDGSLATGGDRSLLDVDRTPKTPVLHSPFACHRDVWTGCATWKDNLISCSGSGSV 2130
1068 QY KWNIIITGNKEKDFVCHQGTVL-----SCDI 1093

2131 DB GLWDPESQRLUGQFLGHQSAVSAVAEEHVSVSRDGTCLKWDHQGVELTSIPAHSGP 2190
1094 QY SHDATKFSSTSA-----DKTAKIWFDFLLPLHLHELGRHNGCVRCSAFSVD 1139
2191 DB SHCAAMEPRAAGQSGSELLVVTVGLDGAATLWHLPLLVCOHTLLGHSGVPVRAAVSETS 2250
1140 QY TLLATGDDNGEIRIWNVNSGELLHLHLCAPLSBEGAATHGWTDLCLFSPDGKMLIS---AG 1196
2251 DB GLMLTASDGSVRLVQVPK-EADDTCTIPRSSAA-----VTAVAWAPDGSMAVSGNOAG 2302
1197 QY GYIKWKN---VVTGE-----SSQTEY 1214
2303 DB ELILWQEKAVATAQAQPGHIGALLIWSSAHTFF 2334

RESULT 10

US-09-184-445-3
; Sequence 3, Application US/091844445
; Patent No. 6174703
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 6174703el Genes Encoding Telomerase Protein
; TITLE OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/184,445
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/751,189
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2627 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-184-445-3

Query Match 6.2%; Score 408; DB 3; Length 2627;
Best Local Similarity 21.0%; Pred. No. 4.1e-27;
Matches 296; Conservative 170; Mismatches 492; Indels 454; Gaps 58;
QY 97 SSGKSDSVSGTISYVRTVLCE-----GGVPQRPVVFVRKKLVNAIQKLSK-- 142
DB 1083 AAGRPYVGLLEFQGLVLDVNNMIQKLYLQPGALLEQ-VSIPDDDLVQATFQLOKPP 1141
QY 143 -----LKPEPGWVTHGMAGCKSVLAABAVRDHSLLEGCPFGV--- 182
DB 1142 SPARPRLLQDTVQXLMPLPHGRSLVLT--GSGGQKTAFLASLV---SALQA--PDGAKVA 1194
QY 183 -----HWVSVGKQKSGLLMKLQNLCTRLDQDESFQRLPLNIEBAKDRILMLRKHPR 237
DB 1195 XLVPFHP-SGARPQGLALTLLRLCTYLRGOLKEPFGALPSTYRSLVWELQORLLPKSAE 1253

Qy 418 FVNSLLFCDRNG-----KSFYYLHDLQVDFLTKNCSQLODLHKKIITQFQ 465
Db 1463 -----ASHSGNPFPLCPAYIVQSLRLLGEGFVERGARCLSDGPIRTTIKRYG 1514
Qy 466 R-----YHOPHTLSPQEDCMYWNF-----LAYHMASAKHKLCA 502
Db 1515 KRLGLEKTAHVLIAAHLWKTCDDPDSG-----TFRSCPPEALKDPLPHLLQSGNHGLLAE 1569
Qy 503 LMFSLDWIKAKTELVG-----PAHLIH--EFVEYRHILDEKDCAVSFNFOEFLSLNGHL 554
Db 1570 FLTNLHVAAYLE-VGLVPDLEAHVLYASSKEANOKLPAADVAV---PHTFLRQOASL 1625
Qy 555 LGRPPFNIVGLGCEPSTSEVQOAKLOAKQVNDGMLYLEWINK-----KNITNLSRLV 610
Db 1626 LTQYPL-LLLQAAOAPESPVCQAPLLTQRWHD--QFTLKWINKPQLKGQOOLS-LT 1681
Qy 611 VRPHTDANYHACFSDGRIASCADXTLOVFAETGEX----- 649
Db 1682 MSSSPTRV---AFSPNGORAAVGTAGTITTYLLNKLWQEEKAVSGDGISSFAFLSDTA 1738
Qy 650 -----LLEIKAHEDVLCAPSTDDRFIATCSVDKVKVWNSMTG 689
Db 1739 LFTTDPGHLEMDLQHCWVFOKAHOYQITCCCLSPDRRLATVCLGGYKLMDTVRG 1798
Qy 690 ELVHTYDEHSEQVNCFTNSSHLLATGSDCFCLKLMDLNQKECRNTPMFGHTNSVNH 749
Db 1799 QLAFOY-THPKSLNCVAFHPBQ--VVATGSWAGSITFPQADGLKVTYKELGAPGPSVCSL 1855
Qy 750 RFPDDKLLASCADGTLKLDATSNANERKSINVQO-----PFLNLED-----POEDME 798
Db 1856 AFNPKGKIVAVGRIDGTVELW-AWQEGARLAAPPAQCGCSAVFLHAGORFLTAGED-- 1912
Qy 799 VIVKCSWSA-----DCARIMVA-AKNKIFLFDIHTSGLLG 833
Db 1913 --GKAQMSGFLGRPGCLGSLPLSPALSVALNPDQGVAVGREYDGINIVKI--SSGSQ 1969
Qy 834 EIHGTGHSTIOYCDFSPONHVALVSQYCVLWNTDSKSVADCRGHLSVWGHWFSPD 893
Db 1970 PQH-----QELNAVSAU-----VLSPSVLYSGAEDGSL---HGMFKGD 2007
Qy 894 GSSFL-----TSDDQTIRLWETKKVKCNKSAVMLKQEVVDVVFQE 932
Db 2008 SLHSLMLSRQKPVGLAASRELAASDEFVRLMFRQ----- 2047
Qy 933 NEWVLAVDHIHRLQJ-----INGRTQIDYLTQAVSCCLSPHLQYIAPG--DENGAI 985
Db 2048 ----LLTQPHVHVELPCCAEIRG-----EGPVCCSFSPDGGILATAGRDRLLC 2095
Qy 986 EILELVNNRIFOSRQ--HKTVVHIOFTADEKTLISSDDABIQVN-----WOLDKCIF 1039
Db 2096 WDKTIAQAPLLIHTFSSCHRWITGCAWTKD-NILVSCSDSGSVGLMNPAGQOOLQO--- 2151
Qy 1040 LRGHQETVDFRLLKNRSLWSFDTGVKVVNNIITGNKEKDFVCHQGTVLSCDISHDATK 1099
Db 2152 FSGHQSAVASV-AVAEEHIVSVSRDGTCLKWD-----HOGVELTSIPAHSGPI 2198
Qy 1100 FSSTSA-----DKTAKIWSFDLLPLHLRGNHGNVCRCSAFSDVS 1139
Db 2199 SQCAAALEPRGCGPGBSELLVTVGLDGAATKLMHPLLVQCINTLQHGSGPVTAASAEAS 2258
Qy 1140 TLLATGDDNGEIRIWNVSGELLHLCAPLSEGAATHGGWVTDLCFSPDGMILISA--GG 1197
Db 2259 GLLLTSDDS-SVOLWQIFK-EADSYKPRSSVA-----ITAVAWAPDGSVMVSGNEAG 2309
Qy 1198 YIKWNN-----VVTGE-----SSOTFYNTGNTN 1219
Db 2310 ELTLWQOAKAVATQAQORVSHLIWYSANGSFFVLISAN 2346

RESULT 13
US-09-184-445-4
; Sequence 4, Application US/09184445
; Patent No. 6174703

GENERAL INFORMATION:
APPLICANT: Harrington, Lea A.
APPLICANT: Robinson, Murray O.
TITLE OF INVENTION: No. 6174703el Genes Encoding Telomerase Protein
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSER: Angen, Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/184,445
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/751,189
FILING DATE: 15-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy A.
REGISTRATION NUMBER: 34,688
REFERENCE/DOCKET NUMBER: A-433
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2629 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-184-445-4
Query Match 6.08; Score 399.5; DB 3; Length 2629;
Best Local Similarity 21.48; Pred. No. 2.5e-26;
Matches 290; Conservative 178; Mismatches 470; Indels 419; Gaps 62;
Qy 120 PORPVVFTKLVNALQKLSKLGEGPWTHMGACGK-----SVLAABAVRDHSL 174
Db 1152 PNP-----RLQDTVQQLLPHGRISLVT--QACQGTAFSLASVSAKVDPQ--- 1199
Qy 175 EECFPGVHWV-----SVGQDKSGLLMKLNCTRLDQDESFSORLPNTEAKDRLIL 230
Db 1200 ----PNEPFFVFFHFAARPDQCLANLLRRLCTHLRQKLGELSALPSTYRGLVWELQK 1255
Qy 231 MLRKHPRSL-----LILD-----DVWDSWLKAFDSQCQILLLTRDKSVTDSV 273
Db 1256 LLLKFAQLQPAQTLVLIIDGADKLVDNRNGQLISDWIPKSLPRVHLVSVS---SDSG 1311
Qy 274 MGP-----KYVPVSESLGKEG---LEILSL-----FVNMKKADLPQAHSHI 314
Db 1312 LGETLQSQGAVVVALGSLVPSSRAQLVREELAYGKREESPFNNQMLLAKQSSSL- 1370
Qy 315 KECKSGPLVWSLIGALIRDF-----PNRWEYVKOLQNKQPKRIRKSS 357
Db 1371 ----PVLHLVTDYLRFTLYEQVSRRLRTPATLPLLIQLHILSTLEO----- 1415
Qy 358 SYDVEALDEAMISVEMLRDIEDIKVYVDLSILQDKVKVPTKVLCLWDMETEVEEDLOE 417
Db 1416 -HGHVLPQALT-ALEVTR-----SGITVDQLHAULST---WLILPKETKSWEVLA- 1462
Qy 418 FVNSLLFCDRNG-----KSFYYLHDLQVDFLTKNCSQLODLHKKIITQFQ 465
Db 1463 -----ASHSGNPFPLCPAYIVQSLRLLGEGFVERGARCLSDGPIRTTIKRYG 1514
Qy 466 R-----YHOPHTLSPQEDCMYWNF-----LAYHMASAKHKLCA 502
Db 1515 KRLGLEKTAHVLIAAHLWKTCDDPDSG-----TFRSCPPEALKDPLPHLLQSGNHGLLAE 1569


```

Qy 503 LMFSLDWTAKTELVG-----PAHLIH--BEVEYRHILDEKOCVAVSNFOEFISLNGHL 554
Db 1570 FLTNLHVVAAYLE-VGLVPDLLEARLVYASSKEANQKLPAAVAV-----PHTFLRQQASL 1625
Qy 555 LGRPPFNIVOLGLCEPETSEYQOAKLOAQEVDNGMLYLEWINK-----KNITNLSRLV 610
Db 1626 LTQYPL-LLLQOAAASQPESVCCOAPLITRWHHD--OFTLKWINKPOTLKGQOQSL-LT 1681
Qy 611 VRPHTDAYHACFSEDGQRIASCADKTLQVFKAEKTEK----- 649
Db 1682 MSSSPTAV---AFSPNGQRAAVGTASGTIYLLNLKTQWEERKAVVSGCGOISSFAFLSDTA 1738
Qy 650 -----LLEIKAEDEVLCAPSTDDRIATCSVDKVKVKNWSMTG 689
Db 1739 LFLATFDGHLBELDQHCWVQTKAQYITGCLSPDRLLATVCLGGYKLKLDWTVRG 1798
Qy 690 ELVHTYDHSQVNCCHFTNSHHLLLATGSSDCFLKLDLWLNQECRTMTEGHTNSVNH 749
Db 1799 QLAFOY-THPKSLNCVAFHPGQ--VWATGSMAGSITFFQADGLKVKELGAPGFSVCSL 1855
Qy 750 RFSDDKLLASCADGTLKLDATSANERKSIINVKQ-----PFLNLED-----POEDME 798
Db 1856 AFNPKGIKIVAGRIDGTVELW-AWQEGARLAAFAQCCCVSAVLFLHAGDRFLTAGED-- 1912
Qy 799 VIVKCCWSA-----DGARIMVA-AKNKIFLFDIHTSGLLG 833
Db 1913 -GRAQLWSGFLGRPGCLGSLPALSPALNPDGQVAVGYREDGINIYKI--SSGSQ 1969
Qy 834 EIHTGHSTIOYCDPSPONHLAVVALSQYVELWNTDSRSKVAOCRHLSHVHGMFSPD 893
Db 1970 PQH-----QELNAVAVSAL-----VMLSPVLVSGAEDGSL-----HGMFPGD 2007
Qy 894 GSSFL-----TSSDDOTIRLWETKVKCKNSAVMLKQEVVDVVFQE 932
Db 2008 SLHSLWLLSRQKVPVGLAASRELMAASEFTVRLWPRQ----- 2047
Qy 933 NEVMVLAVDHIRRLQ-----INGRTQIDYLTBAQVSCCLSPHLQVIAFG--DENGAI 985
Db 2048 ----LLTQPHVAVELPCCAEIRGH-----EGPVCCCSFSPDGGILATAGDRNLLC 2095
Qy 986 EILELVNRRIFQSRFQ--HKTIVWHIQTAEKTLISSDDAEIOVWV-----WOLDKCIF 1039
Db 2096 WDMKIAQAPLIHTFSSCHRWITGCATKD-NILVSCSDSGSVGLMNPAGQQLGQ--- 2151
Qy 1040 LRHQETVKDFLLKNSRLLSWSPDGTVKVWNIIITGNKEKDFVCHQGTVLSCDISHDATK 1099
Db 2152 FSGHQSAVASV-VAVEHIVSVSRDGLKWV-----HOGVELTSIPAHSGPI 2198
Qy 1100 FSSTS-----DKTAKIWSFDLLPLHLRHNCGVRCRCSAFSVDS 1139
Db 2199 SOCAAALEPRPGQPGQSELLVTVGLDGTATKLHPLLVQCIRTQLQHSRGPVTAASAEAS 2258
Qy 1140 TLLATGDNDGIRIWNVNSGELLHLHLCAPLSEGAHGGWTDLCFSPDGKMLISA--GG 1197
Db 2259 GLLLTSDS-SVOLWQIPK-EADDSYKPRSSVA-----ITAVAWAPDGMVSVGNEAG 2309
Qy 1198 YIKWVN-----VVTGE-----SSQTFVTNGTN 1219
Db 2310 ELTLWQQAQAVATAQAQGRVSHLIWYSANSFVL SAN 2346

```

```

RESULT 14
US-09-291-170A-13
; Sequence 13, Application US/09291170A
; Patent No. 6410687
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald D.
; APPLICANT: Hartman, James J.
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; FILE REFERENCE: 18557B-000510US

```

```

; CURRENT APPLICATION NUMBER: US/09/291,170A
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 60/081,734
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 251
; TYPE: PR1
; ORGANISM: Thermomonospora curvata
; FEATURE:
; OTHER INFORMATION: putative serine/threonine kinase PKwa WD40 repeat
; OTHER INFORMATION: region
; US-09-291-170A-13

Query Match 4.9%; Score 322; DB 4; Length 251;
Best Local Similarity 28.1%; Pred. No. 5.9e-21;
Matches 84; Conservative 47; Mismatches 104; Indels 64; Gaps 7;

Qy 614 HTDAYHACFSEDGQRIASCADKTLQVFKAEKTEKLEIKAEHDEVLCAPSTDDRI 673
Db 12 HTDWVRAVAFSPDGCALLASGSDATVRLWDVAAAEERAVFEGHTHYVLDIAFSPD 71
Qy 674 TCSVDKVKVKNWSMTGELVHTYDEHSEQVNCCHFTNSHHLLLATGSSDCFLKLD 733
Db 72 SGRDGTARLWN-----VATGTEHAVLK----- 94
Qy 734 ECRNTEGHTNSVNHCRFSPPDKLLASCADGTLKLDATSANERKSIINVKQFFLNLED 793
Db 95 -----GHTDYVAVAFSPDGSVMASGRDGTIRLWDVATGKERDV-----LOAP 138
Qy 794 QEDMEVIVKCCWSADGARIMVAANKIKIFLDIHTSGLLGE-IHT--GHHSTIQVCD 850
Db 139 AEN-----VSLAFSPDGSMLVHSGDSTVHLWDVAS-----GEALHTFEGHTD 190
Qy 851 QNHLLAVVALSQYVELWNTDSRSKVAOCRHLSHVHGMFSPDSSFLTSSDDDTIRL 909
Db 191 DGALLASGSDRTIRLWDVAAQEBHTTEGTEPVSVAHPHEGTTLASASEDGTIRW 249

RESULT 15
US-09-724-884-13
; Sequence 13, Application US/09724884
; Patent No. 6429304
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald D.
; APPLICANT: Hartman, James J.
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; FILE REFERENCE: 18557B-000510US
; CURRENT APPLICATION NUMBER: US/09/724,884
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/291,170
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 251
; TYPE: PR1
; ORGANISM: Thermomonospora curvata
; FEATURE:
; OTHER INFORMATION: putative serine/threonine kinase PKwa WD40 repeat
; OTHER INFORMATION: region
; US-09-724-884-13

Query Match 4.9%; Score 322; DB 4; Length 251;
Best Local Similarity 28.1%; Pred. No. 5.9e-21;
Matches 84; Conservative 47; Mismatches 104; Indels 64; Gaps 7;

Qy 614 HTDAYHACFSEDGQRIASCADKTLQVFKAEKTEKLEIKAEHDEVLCAPSTDDRI 673
Db 12 HTDWVRAVAFSPDGCALLASGSDATVRLWDVAAAEERAVFEGHTHYVLDIAFSPD 71

```

